SPRINGTIME ONSET OF ISOLATED CONVECTION IN THE

CENTRAL AND EASTERN UNITED STATES

by

Christopher M. Jarrett

July, 2021

Director of Thesis: Thomas M. Rickenbach, Ph.D.

Major Department: Geography, Planning, and Environment

In the southeast United States (SE US), precipitation is present year-round though with a variety of distinct regimes. Convection organized as isolated precipitation features (IPF), typically isolated afternoon thunderstorms, has a strong annual cycle and accounts for nearly 40% of the total rain during summer. When extracted from the annual precipitation cycle, the seasonal evolution of IPF rain in the SE US is reminiscent of the summer monsoon in other parts of the world. Close examination of this hidden monsoon-like regime of summer thunderstorms may guide our understanding of how subtropical precipitation regimes change as the Earth warms. Recent studies for the SE US have shown that onset of the IPF summer precipitation regime is abrupt, occurring over a few weeks in May. The spatial range of IPF onset across the US is not known. This work (1) extends the analysis to a nine-year temporal period (2003-2011), (2) expands the spatial extent into the Central Plains and northeast US, and (3) determines onset within an array of box regions of 4^o x 4^o and 2^o x 2^o in size, while using a new hourly radar dataset. The impact of different array size and rain threshold on the timing and geographic pattern of onset was examined. Results show that for the nine-year average, IPF onset is generally centered geographically on the SE US, but also extends into the mid-Atlantic region

and the lower Mississippi River valley. The 4^{o} x 4^{o} box array using the climatological average IPF rain for the SE US as a threshold performed well for the IPF onset analysis. On average, IPF onset begins at the end of April (pentad 24) in south Florida and continues across the southeastern coastal plain into early and mid-May (pentads 27-28). Onset progresses northward and westward in early June (pentads 31-32), with the latest onset in eastern Ohio and the Texas Gulf Coast by mid-late June (pentads 35-36). These results support the hypothesis that the springtime increase in moisture and instability from the seasonal establishment of the North Atlantic subtropical high is mainly limited to the SE US and is responsible for conditions leading to IPF onset.

SPRINGTIME ONSET OF ISOLATED CONVECTION IN THE CENTRAL AND EASTERN UNITED STATES

A Thesis

Presented To the Faculty of the Department of Geography, Planning, and Environment

East Carolina University

In Partial Fulfillment of the Requirements for the Degree

M.S. Geography

by

Christopher M. Jarrett

July, 2021

SPRINGTIME ONSET OF ISOLATED CONVECTION IN THE

CENTRAL AND EASTERN UNITED STATES

by

Christopher M. Jarrett

APPROVED BY:

DIRECTOR OF
THESIS: _____
(Thomas M. Rickenbach, Ph.D.)

COMMITTEE MEMBER: _____
(Rosana Nieto Ferreira, Ph.D.)

COMMITTEE MEMBER: _____
(Hannah Cooper, Ph.D.)

CHAIR OF THE DEPARTMENT
OF GEOGRAPHY, PLANNING, AND
ENVIRONMENT:_____
(Jeff Popke, Ph.D.)

DEAN OF THE
GRADUATE SCHOOL: _____
Paul J. Gemperline, PhD

ACKNOWLEDGEMENTS

This study was funded in part by a grant (AGS-1660049) from the National Science Foundation Climate & Large-Scale Dynamics Program.

My time at East Carolina University has been nothing short of a roller coaster ride experience. I faced many trials both academically and personally, and it was because of this that I grew as a student, scientist, and man. These accomplishments are not a product of my own efforts exclusively, as there were many people who made this journey and its successes possible. This journey dates all the way back to my undergraduate days at ECU. I would like to thank the professors from both the Department of Health and Human Performance and the Department of Geography, Planning and Environment, as they were the ones who made graduate school possible in the first place. I also send thanks to my family for their unwavering support, even if they don't understand what I am working on or how it will better my future. I would also like to extend thanks to my advisor of six years extending back all the way to my sophomore year of undergraduate studies, Dr. Tom Rickenbach, and my committee members, Dr. Rosana Nieto-Ferreira and Dr. Hannah Cooper. Their advice, encouragement, expertise, and mentorship have all played huge roles in my growth as a scientist and as a researcher throughout this study. I also would like to thank my friends, both near and far, and my girlfriend for their unconditional support through both the good and the bad. I would also like to extend thanks to my fellow graduate colleagues in the Department of Geography, Planning and Environment. They helped make this experience quite enjoyable, and I am forever grateful for the friendships that we built over the last two years. Finally, I would like to thank my dog, Cosmo, for always giving me a reason to smile when returning home and for making graduate school a bit easier.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Precipitation regimes vary vastly across the continental United States from season-to-season. The west coast experiences dry summers and wet winters due to the cold season southward displacement of the Pacific subtropical high (Markham, 1970). The intermountain west and Rocky Mountains experience winter snow and orographically-forced thunderstorms in the summer. Southerly low-level jets associated with extratropical cyclones (Djurić and Ladwig, 1983) are responsible for winter season precipitation in the Great Plains, while mesoscale convective systems are the primary precipitation producer during the warm season (Carbone and Tuttle, 2008). The southeast (SE) United States winter is characterized by a mix of rain and snow from extratropical cyclones. Summer precipitation is dominated by diurnal thunderstorms associated with modulation of the North Atlantic Subtropical High and reversal of coastal sea breezes (Prat and Nelson, 2014). In the fall season, a significant amount of precipitation comes from tropical cyclones in the SE US (Prat and Nelson, 2013). Winter precipitation in the northeast US come in a variety of forms, including Nor'easters and lake effect snow. Frontal systems and northward tracked mesoscale convective systems account for much of the northeast US summer precipitation, with a limited contribution from tropical cyclones.

It is useful to study the seasonal changes in precipitation regime for different geographical regions of the US to better understand the long-term variation of precipitation linked to periodic drought, ENSO, and human-induced climate change. In particular, the SE US receives ample precipitation on a year-round basis (Rickenbach et al., 2015). By contrast, the central United States exhibits a clear summer maximum in average precipitation (Prat and Nelson, 2015). Studies related to monsoon systems offer a framework that is suitable for examining seasonal changes in precipitation. Using four years of radar-based precipitation data,

Rickenbach et al. (2020) studied the spring-to-summer onset of the locally generated isolated thunderstorm regime in the SE US for a four-year period by applying a methodology that was used in determining the timing of monsoon onset in South America. Their study left unanswered the question of the geographic extent of the seasonal isolated thunderstorm domain beyond the SE US.

This thesis will address that question by broadening the Rickenbach et al. (2020) analysis in time and space. Using an expanded radar-based precipitation data set, the current study will analyze the spring to summer onset of the isolated thunderstorm regime for a seven-year period based between 2003 and 2011 across the entire eastern half of the United States. Geographic expansion of this analysis will include the central and northeast US regions, in addition to performing a reanalysis of the SE US over the temporal period. The onset methodology and its sensitivity will be tested with respect to the spatial dimensions of the local analysis regions. The application of this onset methodology will be used to construct precipitation time series and determine onset of the isolated convective season across the US. Regional differences in onset will be studied by direct comparison of each of the three regions against one another. Further, the nature of onset for each region will be studied individually on an year-to-year temporal scale. The following questions will be explored:

1. How does the springtime onset of the isolated convection season differ temporally and spatially across the United States?
2. What is the spatial extent of the seasonal isolated convection regime in the eastern United States?
3. How does the temporal and spatial nature of onset vary year-to-year within each individual region?

4. How sensitive is the date of onset to changes in the size of local analysis regions and onset threshold?
5. Are there spatial patterns of onset timing that are correlated to geographic features such as orography, coastal regions, or land surface type?

CHAPTER 2: LITERATURE REVIEW

This chapter presents a review of literature related to the topics of convection, seasonal changes in convection in the SE US and central US, year-to-year variability of rainfall in the SE US and central US, and the general concepts related to monsoon onset. This section starts with an overview of basic concepts related to the formation of isolated and organized deep convection, to provide a basis of understanding for the sections to follow. Following this is a section on mesoscale convection and various sources of convection to introduce a hierarchy of convective organization. The third and fourth sections alike discuss rainfall variability for the SE US and central US and related convective factors. Finally, the fifth section reviews work related to monsoon climates and methods of identifying the timing of monsoon onset as a basis for identifying the isolated convective season onset in this project.

2.1 Concepts Related to Isolated and Organized Deep Convection

To understand convection and its role in large-scale precipitating systems, it is important to understand the key factors and concepts that have an influential role in convective processes. Important to the processes of convection, cloud growth, and storm production are the conditions of tropospheric stability, the mechanisms that allow air to rise, and atmospheric controls on convection.

2.1.1 Tropospheric Stability and Buoyancy

The stability of the troposphere is one of the most important factors for convective initiation and storm development. Static stability is defined as the resistive force offered to various motions by atmospheric structures of vertical density (Gates, 1960). As such, the state of stability in the troposphere acts as a control for the inhibition or allowing of vertical air ascent. When the troposphere is statically stable, air cannot easily rise. Important to stability is the force

known as buoyancy. Buoyancy is the upward force exerted on a parcel in a gravitational field as a result of differences in parcel and environmental densities (Doswell III & Markowski, 2004). Positive buoyancy simply means that a warmer, less dense parcel will lift when surrounded by a colder, denser environment. Conversely, negative buoyancy means a cold, denser parcel will sink when surrounded by warm, less dense air. Under the influence of a stable troposphere aloft, the environmental temperature is typically warmer than a parcel of air trying to ascend. In this case the buoyancy force is negative, and the air parcel will move back towards its initial position (Peppier, 1988).

The most common state of the troposphere is conditionally unstable. In a state of conditional instability, the atmosphere below a cloud base is stable, while the region encompassing the cloud itself is unstable. Because the area below the cloud is stable and the air is unsaturated, it cannot rise on its own, and must therefore be forced upward. Once saturated, the parcel is considered unstable and will rise without similar forcing (Peppier, 1988). However, instability can only lead to convection if air is forced to ascend to the LFC, or the height in the troposphere in which an air parcel raised moist adiabatically becomes warmer than the environment surrounding it and now rises freely due to a positive buoyancy force.

2.1.2 Mechanisms of Forced Ascent

When air is unsaturated it requires forcing to initiate lifting below the LFC. One of the primary means of air lifting below the LFC is that due to forced frontal ascent. Forced frontal ascent occurs when a region of warm air converges with a cooler and denser front, causing the warm and less dense air to rise. This surface convergence occurs at frontal boundaries, enhanced by upper troposphere divergence ahead of the jet stream trough. This couplet of surface convergence and upper-level divergence results in the formation of a surface low, ahead of the

jet stream trough. Orographic lifting is a second source of forced air ascent. Air that blows against the windward side of mountains ascends along the upsloping terrain. This lifted air often reaches saturation, which increases hydrometeor content of clouds and acts as a means of initiating moist instability, especially given the presence of differential heating over the sloping terrain (Kirshbaum et al., 2018). Thunderstorm outflow boundaries, or gust fronts, also act as a mechanism of ascent in the boundary layer. Koch and Ray (1997) state that gust fronts are the leading edge of downdrafts that are caused by evaporative cooling and precipitation drag. These downdrafts are cooler than the low-level winds, so the warmer air is forced upward if it is moist enough. As such, it becomes possible for new convective cells to be generated.

Of great importance to warm season convection are sources of convective boundaries and convergence zones in both the SE US and central US. Using observations from WSR-88D radars, Koch and Ray (1997) defined a boundary as a region of high reflectivity with a width of 10 km or less with a lifespan of greater than 1 hour. While the study area was central and eastern North Carolina, many of the observed boundaries are relevant for the SE US, including thunderstorm outflow boundaries, sea-breeze fronts, synoptic scale fronts, shallow fronts and prefrontal troughs, all of which may act as forces of ascent (Koch and Ray, 1997). Combined with the results from Wilson and Schreiber on a similar study based in Colorado (1986), topographically induced boundaries, like the sea breeze or convergence lines east of the Rockies, produce convection about 80% of the time in both North Carolina and Colorado (Figure 1) (Koch and Ray, 1997). However, frontal boundaries and thunderstorm outflows produce convection far more in North Carolina than in Colorado due to differences in moisture in the planetary boundary layer and resulting CAPE. This does not come as a surprise, given that North Carolina is situated along the east coast of the US and therefore has more moisture available due to their proximity to the Atlantic Ocean.

2.1.3 Convective Controls

2.1.3.1 Convective Available Potential Energy

Houze (1993) states that convective available potential energy (CAPE) is a representation of environmental buoyant stability and is defined by Barkiđija and Fuchs (2013) as,

$$CAPE = g \int_{LFC}^{EL} \left(\frac{T_p - T_e}{T_e} \right) dz,$$

where T_p is parcel temperature, T_e is environmental temperature, and LFC and EL are the heights of the Level of Free Convection (the height where the buoyancy force is first positive) and Equilibrium Level (the height where the buoyancy force switches from positive to negative), respectively. From this definition, CAPE is the sum of positive buoyancy above the LFC that represents the energy available for air parcel ascent until stability is reestablished at the EL. CAPE is generated by the processes of heat and moisture transport to the planetary boundary layer and temperature and moisture advection (Moncrieff, 2010). To this point, more CAPE will be available as temperature and moisture both increases. When looking at CAPE, the typical threshold value indicative of moderate instability is 1000 J kg^{-1} (DeRubertis, 2006). However, it is possible for convection to form without meeting this threshold. Sherburn and Parker (2014)

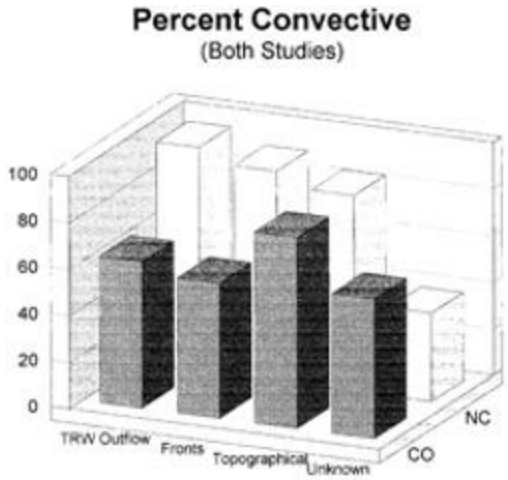


Figure 1: Boundary types and the percentage of time in which convection was produced. From Koch and Ray, 1997.

created a climatology of significant severe convection for environments under the influence of high wind shear ($\geqslant$ 18 m s^{-1} at 0-6 km) and low CAPE ($\leqslant$ 500 J kg^{-1}). While these situations are not as frequent in occurrence, they are still of interest for severe weather climatologies.

2.1.3.2 Convective Inhibition

Just as CAPE is the sum of positive buoyancy above the LFC, convective inhibition (CIN) is the sum of negative buoyancy below the LFC (Moncrieff, 2010). Barkidija and Fuchs (2013) define CIN as,

$$CIN = g \int_{z_{bottom}}^{z_{top}} \left(\frac{T_p - T_e}{T_e} \right) dz,$$

where T_p is parcel temperature, T_e is environmental temperature, and z_{top} and z_{bottom} are upper and lower height bounds. A good way to think of CIN is as negative CAPE. Whereas CAPE is a measure of instability above the LFC, CIN is the exact opposite. As given by its name, CIN's ultimate purpose is to inhibit the startup of convection and instability by attempting to stop a parcel from rising in the stable layer below the LFC. CIN also exists in the region of the troposphere above the EL, where a parcel becomes stable again after its temperature becomes cooler than that of the environment. In general, CIN increases in the lowest heights of the troposphere at night as the surface cools and erodes as surface heating occurs during the day. The erosion of CIN due to daytime surface heating is the reason why warm season convection is most prominent in the afternoon hours.

2.1.3.3 Vertical Wind Shear

While CAPE is a generally good indicator of convection and convective development, it is not a good indicator of convective organization. The presence of wind shear, however, is a major player in the type of convective organization that is present. Typically, sheared

environments are more typical of organized convection, like Mesoscale Convective Systems (MCS), while unorganized convection (isolated storm systems) can be associated with unsheared environments. Holton (1993) defines MCS as cloud systems occurring in connection with thunderstorm ensembles, producing a precipitation area at least 100 km in horizontal scale in at least one direction. Holton (2004) classifies convective storm type based on wind shear speed. Single cell storms tend to have shear less than 10 m s^{-1} below 4 km in height. Multicell storms tend to form under moderate wind shear (10-20 m s^{-1} below 4 km). Finally, large shear environments (greater than 20 m s^{-1} below 4 km) tend to tilt convective cells and produce a cell with an updraft core and trailing downdrafts (Holton, 2004). The multicell storms can form larger-scale squall line storms that combine multiple individual cells together. This is because moderate wind shear helps keep the gust front near storm updrafts which enables the regeneration of convection and the creation of new cells, ultimately assisting with the creation of the squall line (Koch and Ray, 1997).

2.2 Mesoscale Convection and Sources of Convection

Convective organization is not a simple process, as there are many factors and controls that determine organization. In a discussion of seasonal convective diurnal cycles, Wallace (1975) states that two general mechanisms are important: thermodynamic processes that affect static stability and dynamical processes that influence planetary boundary layer convergence. Moncrieff (2010) states that convective organization involves dynamical processes like wind shear, interaction between convection and waves, and maintenance of atmospheric circulations to prevent dissipation. In characterizing various types of isolated thunderstorms, Houze (1993) discusses that both wind shear and static stability both influence the type of storm that may be present. One of the most important factors necessary for convective organization, and convective

initiation in general, is the presence of vertical air motion from thermodynamic buoyancy (Houze, 2004). The issue, however, is that not all storms are organized in the same way.

A hierarchy of organization is a good way to look at convective organization because of its progression in both scale and degree of organization. This hierarchy contains the following kinds of systems: isolated thunderstorms (Houze, 1993), mesoscale convective systems (MCS) (Leary and Houze, 1979, Houze, 1989, Houze, 2004), and tropical cyclones (TC).

2.2.1 Isolated Thunderstorms

Isolated thunderstorms, where convection is often driven by thermodynamic instability, can be broken down into three classifications: single cell, multicell, and supercell (Houze, 1993). The single cell classification is a cumulonimbus cloud with one precipitation band. These clouds typically are composed of ice in its upper region and often spread out to form the classic anvil shape. On the developing side of cumulonimbus clouds are vertical updrafts and corresponding updraft cores and cloud turrets (Houze, 1993). The updrafts of greater velocity also favor high condensation rates and corresponding graupel production from supercooled water collected by ice particles, some of which are carried outward to form the anvil before melting and falling as rain. Of great importance are the presence of downdrafts that form as a result of evaporative cooling (Wallace, 1975). These downdrafts can sink as far as the earth's surface and can spread to form a gust front. The gust front enables further lifting to create new clouds, resulting in the formation of multicell thunderstorms. As such, it is important to note that multicell storms consist of cells at different maturation stages. This is important because the stage of cell maturation is indicative of updraft and downdraft behaviors. The cells that are early in maturation contain intense updrafts that allow for hydrometeor growth. Mature cells consist of well-defined updrafts and downdrafts, allowing for heavy precipitation. Finally, cells that are

considered dying only have a downdraft, which will precipitate and eventually dissipate (Houze, 1993). Supercell thunderstorms are typically the same size as multicell thunderstorms but are characterized by a single, large updraft-downdraft pair and are typically associated with having regions of intense cyclonic vorticity (Houze, 1993). While each of these three types of isolated storms have unique characteristics, their size and period of life are much less than that of MCS or tropical cyclones. The lifetime of isolated storms tends to be only a few hours, while their size does not typically exceed the size criterion to be considered mesoscale (Houze, 1989).

2.2.2 Mesoscale Convective Systems

MCS's are the largest form of deep convective storms (Feng et al., 2019) and are related to several mesoscale circulations (Houze, 2004). Their organization is interesting because they often come in many forms. Related to the discussion above on isolated thunderstorms, MCS can form from multiple isolated convective cells that have merged together (Doswell III, 2001). In situations like this, the new "line" of thunderstorms takes on a straight or curved appearance and tend to be longer lived than if the isolated convective cells were separated from one another (Houze, 1993). This is called a squall line, and this long line of merged storms runs perpendicular to the low-level wind direction. MCS in the warm season US tend to form under the influence of weak baroclinic forcing (Feng et al., 2019), though squall line MCS require strong low-level wind shear. Weak baroclinic forcing consists of ridge influence from high pressure and weak low-level convergence acting upon warm, moist air to cause ascent. Further lifting may come as a result of forced ascent along frontal boundaries.

MCS consist of two specific regions: a convective region and a stratiform region (Figure 2). The convective region consists of heavy precipitation with deep convective cores. Conversely, the large stratiform region consists of lighter precipitation and is produced in part

due to the dissipation of older convective cells (Houze, 2004). Within MCSs, upward motion often begins in the boundary layer near the gust front and extends into the convective region. On this upflow in the convective region are localized updrafts and downdrafts of high intensity, resulting in the formation of new convective cells ahead of the heavy rain (Houze, 1993). From this point on, the development of the new cells almost mirrors the development of the multicell thunderstorms discussed previously. The formation of the stratiform region of an MCS takes some time, typically a few hours. During the formative and intensifying stages of the MCS, the stratiform rain area is not present. As precipitation occurs during this time, a low-level cold pool is present. As this cold pool strengthens, a slantwise circulation develops. As new convective cells are superimposed upon this circulation, older cells are advected rearward and the stratiform rain region is created (Houze, 1993).

2.2.3 Tropical Cyclones

Though technically synoptic scale in nature, tropical cyclones are influenced greatly by mesoscale convective processes. As such, tropical cyclones are the third system in the convective organization hierarchy. There are four conditions related to convection that must be met for tropical cyclogenesis. These conditions are as follows: (1) sufficient thermal energy from the

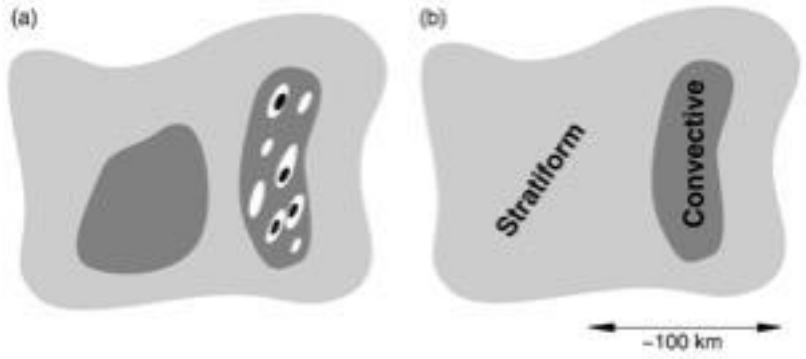


Figure 2: The stratiform and convective regions of an MCS. From Houze, 2004.

ocean, (2) conditional instability, (3) lower troposphere relative vorticity, and (4) weak vertical wind shear. Leary and Ritchie (2010) state that low-level convergence and upper-level divergence, low-level vorticity, and little wind shear over the center are necessary components for tropical cyclone development. The fourth condition relating to wind shear is an interesting one, since thunderstorms, especially of the supercell variety, benefit from the presence of strong vertical wind shear. Shearing processes can tilt the vortex, ultimately distorting the shape and organization of the cyclone. Discussing cumulus convection in hurricanes, Smith (2000) states that convergence above the boundary layer is required for hurricane intensification. This convergence typically comes from latent heat release. Latent heat release, powered by oceanic moisture flux, is therefore one of the major energy sources for tropical cyclone development and is transported upward in tall convective clouds.

Leary and Ritchie (2010), in a study related to the role of convection Eastern North Pacific tropical cyclone genesis, state that storms in the pre-tropical cyclone phase can consist of one or multiple MCS's. These MCS-based setups surround a central point of loosely organized deep convection. Houze et al. (2009) describe the life cycle of an MCS during tropical cyclone development (Figure 3). MCS's begin as one or more rotating, deep convective updrafts, called vortical hot towers (Smith, 2000). These towers that are close to one another can form into an MCS. The stratiform region of these MCS's tend to be associated with a mesoscale convective vortex at the middle heights of the troposphere, which is created from leftover potential vorticity from dissipating convective cells. During tropical cyclone development, the clouds are made up by a combination of both the individual convective towers and maturing MCS's.

2.3 Precipitation Variability in the Southeast United States

In recent years, a lot of work has been done to create a precipitation climatology for the SE US. Rickenbach et al. (2015) created a four-year rainfall climatology for the SE US based on precipitation feature size, to distinguish isolated or multicell thunderstorms from organized mesoscale convection. The purpose of that climatology was to study the seasonal progression of these two general classes of convective organization. More specifically, precipitation features observed by radar were identified as either isolated or mesoscale, depending on whether the feature met a horizontal dimensional threshold of 100 km in length (Rickenbach et al., 2015). Separating precipitation features by size is not a new concept, with some studies dating back as far as that by Leary and Houze (1979). This study looked at six different mesoscale features associated with one tropical cloud cluster off the West African coast. Of these six features, precipitation patterns, lifetime and maximum horizontal dimension were determined manually using four shipboard radars. The importance in being able to identify the size of these precipitation features is because the forcings related to different types of precipitation features differ (Nesbitt et al., 2006, Rickenbach et al. 2020). Features with a horizontal dimension of 100

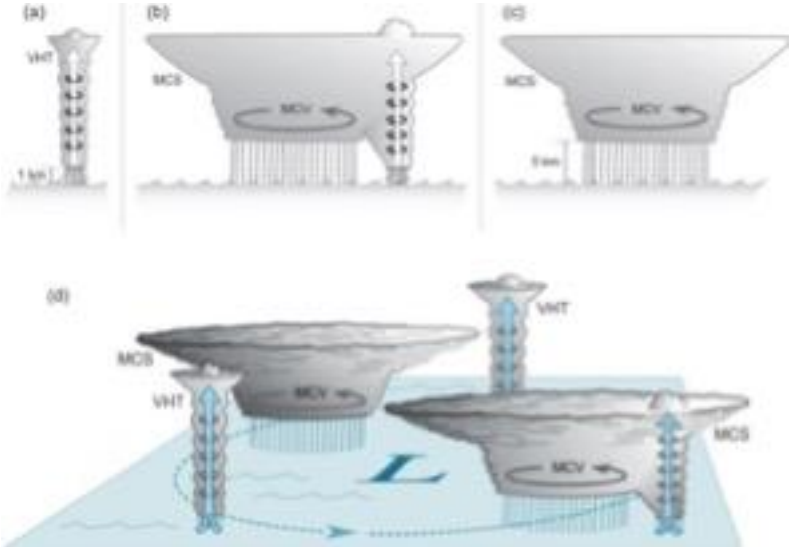


Figure 3: MCS lifecycle during Tropical Cyclone Development. From Houze et al., 2009.

km or greater are classified as mesoscale precipitation features (MPF), while features with horizontal dimension less than 100 km are classified as isolated precipitation features (IPF) (Rickenbach et al., 2015). Formation of the larger-scale MPF precipitation is due primarily to larger-scale synoptic controls, such as forced frontal ascent (Diem, 2006, Rickenbach et al., 2020). IPF precipitation can be attributed to more localized controls, such as the sea breeze, topographic circulations and boundary layer thermal circulations (Carbone and Tuttle, 2008, Rickenbach et al., 2015), and typically include pulse thunderstorms whose convection forms without a synoptic forcing mechanism (Miller and Mote, 2017). In the SE US, MPF precipitation does not have a distinct seasonal cycle. On average, 70% to 90% of yearly precipitation can be attributed to these large-scale systems (Rickenbach et al., 2020). IPF precipitation makes up a smaller portion of average yearly precipitation for the SE US but has a clear seasonal cycle with a warm season maximum. IPF precipitation also has a distinct diurnal cycle that exists in the warm seasons, where precipitation forms offshore during the night and morning hours and flips inland by mid-afternoon (Rickenbach et al. 2015).

Prat and Nelson (2014) created a thirteen-year (1998-2010) precipitation climatology using remotely sensed data from the Tropical Rainfall Measurement Mission (TRMM) Precipitation Radar and Multi-satellite Precipitation Analysis. It was found that seasonal spatial differences in precipitation maximums exists between the cold and warm seasons. In the cold seasons, more intense rainfall tends to fall west of Atlanta, Georgia, due mainly to synoptic-scale precipitation systems coming from the west. For the warm season, maximums in rainfall intensity are found along the east coast as a result of sea breeze effects (Figure 4) (Prat and Nelson 2014). Much like that of Rickenbach et al. (2015), it was found that there were seasonal differences in diurnal cycles. The cold season favored daytime rainfall with no clear land-ocean difference, while the

warm season was found to have a clear diurnal cycle with opposite phases for land and ocean. Specifically, oceanic rainfall on average was present during the hours of 12:00 to 15:00 LST, with a transition to inland rainfall between 15:00 and 18:00 LST associated with enhanced daytime heating over land (Prat and Nelson, 2014). The daytime heating, in addition to the moisture transfer during the diurnal reversal of the sea breeze, primes the inland troposphere for convection and corresponding cloud development and precipitation.

Midlatitude cyclones (MLC) are major players in precipitation variability for the SE US. Ferreira et al. (2013) found that about twice as many MLCs affected North Carolina in the winter as compared to the summer, related to the poleward shift in the jet stream away from the SE US during the summer. Variability in MLC structure also exists on a seasonal basis, with summer season MLCs characterized by northward displaced, slower propagating, semi- stationary midlatitude troughs that are tilted in a northeast-southwest direction with more widespread

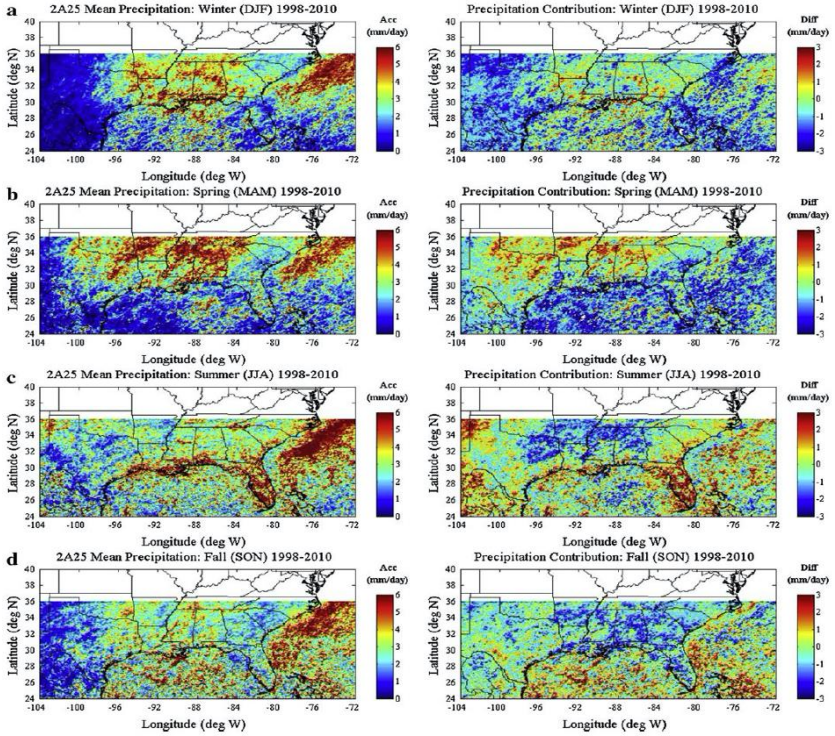


Figure 4: *Mean precipitation and precipitation contribution from the TRMM PrecipitationRadar for each season between 1998 and 2010. From Prat and Nelson, 2014*

precipitation patterns (Ferreira et al., 2013). Compared to 70% of wintertime precipitation in the Carolinas occurring due to MLCs, only about 30% of summertime precipitation can be attributed to MLCs (Figure 5) (Ferreira et al., 2013). Further, MLCs tend to produce more precipitation on the windward side of the Appalachian Mountains compared to the lee side. This finding indicates that MLCs crossing the Appalachian Mountains are subject to a rain shadow effect, meaning that the Carolinas, in particular the central Carolinas, receive less rainfall from MLCs compared to windward regions, such as the Mississippi Valley. On an year-to-year scale, ENSO causes further variability related to MLC precipitation. During the Nino phase, upper-level jets are more intense and southwardly displaced, resulting in more intense precipitation over a larger area compared to the Nina phase.

Another major player in the SE US convective season precipitation variability is the presence of the North Atlantic Subtropical High (NASH). The NASH is a semi-permanent subtropical system of high pressure, or anticyclone, and plays a role in influencing North

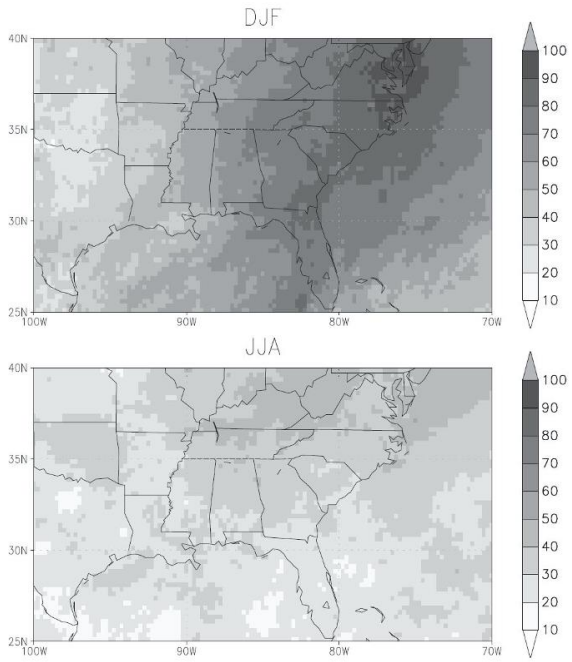


Figure 5: Percentage of monthly rainfall that occurred the day before and the day of MLC passage into North Carolina for winter and summer. From Ferreira et al., 2013.

American weather and climate (Davis et al., 1997). Of importance to this study is the April and May transition from the winter to the summer convective regime in the SE US, which is characterized by westward expansion and strengthening of the NASH in early May (Li et al., 2011). Of particular importance to the SE US is the location of the NASH's western ridge. For the SE US, latitudinal variability of the NASH western ridge is an important factor regarding how much precipitation is observed during a given warm season. When the NASH's western ridge is displaced to the north, it tends to reduce rainfall amounts as the NASH situates itself over the SE US(Nieto-Ferreira and Rickenbach, 2020, Nieto-Ferreira and Rickenbach, 2021). Conversely, the SE US receives enhanced precipitation when the western ridge is displaced southward, as warm moisture from the Gulf of Mexico is transported via winds along the western edge of the high (Li et al., 2011). This warm moisture transport has to do with a sea surface temperature (SST) asymmetry that is common with subtropical highs and is certainly would contribute to the presence of deep convective processes given a proper lifting mechanism. SSTs are warmed on the westward side of the NASH due to poleward ocean currents driven by the southerly wind, while equatorward advection and subsidence cool SSTs on the eastward side of the high. As such, the eastward portion of the NASH experiences suppressed convection and a stabilized atmosphere, while the westward region is characterized by a destabilized atmosphere and deep convection (Seager et al., 2003). A study by Luchetti et al. (2017) explored this relationship between moisture availability and western ridge location as it pertains to North Carolina sea breeze events. Between 2009 and 2012, wet sea breeze events made up 20%-30% of warm season coastal precipitation in North Carolina. These wet sea breeze events are characterized by enhanced values of CAPE and suppressed CIN, as a result of moisture availability due to the moist, warm sea breeze.

2.4 Precipitation Variability in the Central United States

Just as with the SE US, much research has gone into establishing a precipitation climatology for the central US. Using a twelve-year summer-season radar data set, Carbone and Tuttle (2008) found that much of the warm season rainfall in the central US is nocturnal, meaning that much of the rain occurs during the overnight hours and remains present into the morning hours (ex. 0700 UTC) (Figure 6). This nocturnal rainfall occurs frequently, covers a large area and is mainly due to the summer season diurnal cycle of the low-level jet east of the Rockies (Carbone and Tuttle, 2008). The acceleration of the nocturnal low-level jet increases low-level vertical wind shear at night and leads to a favorable environment for nighttime

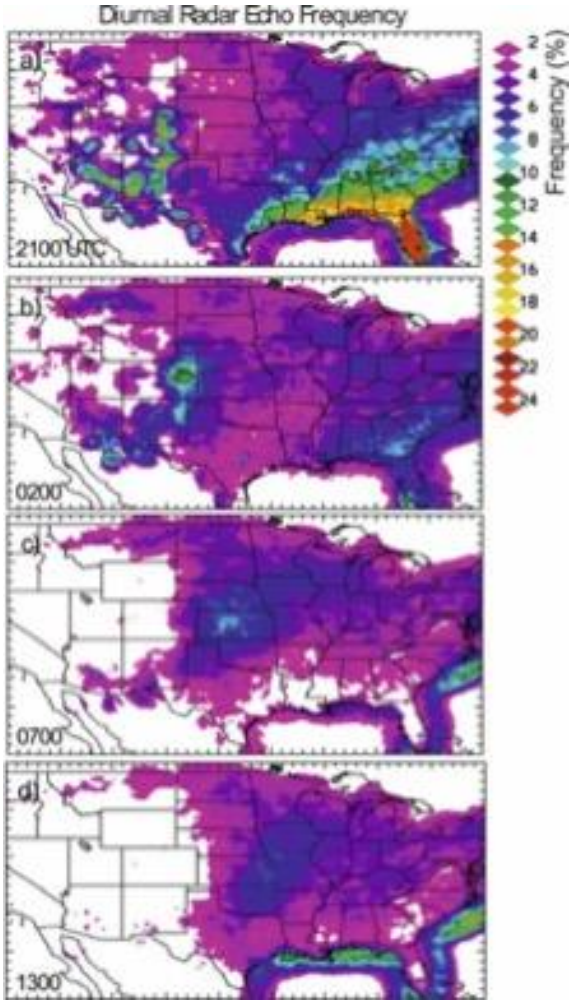


Figure 6: Percent of days within a given hour that a precipitation radar echo is present during the summer season. Plots b and c shows the nocturnal maxima of rainfall over the central US. From Carbone and Tuttle, 2008.

organized convection.

As with the SE US, eastward propagating precipitation systems play into the yearly rainfall in the central US. Of great importance to the central United States is the eastward passage of mesoscale convective systems (MCS). Feng et al. (2019) established a climatology for long-lived (> 6 hours temporal length) MCS for the continental US regions east of the Rockies. Unlike the SE US, where MCSs are present year-round, the central US experiences seasonal variability of MCS occurrence (Figure 7) (Feng et al., 2019). For the central US, the seasonal MCS cycle peaks in the spring period between May and June. The spring and summer season MCSs often consist of deep convection and large stratiform rain regions, which jointly act to increase the magnitude of precipitation (Feng et al., 2019). For the Northern Plains, Klimowski et al. (2003) stated that maximum convective activity during the warm season occurs between 1600 and 1700 LST, in addition to storms preferentially forming along the Rockies during the mid-afternoon portion of the day and maturing into the evening as they propagate eastward. In the Great Plains, an average of 12-16 MCSs occur during the spring and summer seasons, accounting for between 50%-70% of total precipitation (Figure 8) (Feng et al., 2019).

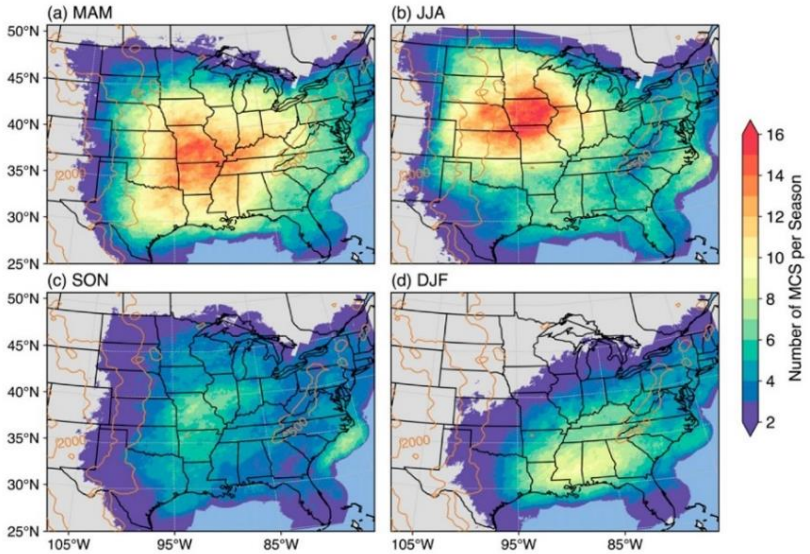


Figure 7: Spatial distribution of MCSs per season from 2004-2016. From Feng et al., 2019.

Haberlie and Ashley (2019) found that most of the MCS-associated rainfall for the warm season tends to occur between the hours of 0400 UTC and 1200 UTC, which is consistent with the findings of Carbone and Tuttle (2008). Consistent with the findings of Feng et al. (2019), Haberlie and Ashley found that more than 50% of total precipitation comes from warm season MCSs in the Central Plains (2019). Unlike the SE US, the central US has less proximity to a source of warm moisture. Whereas the SE US can receive warm, moist air with the sea breeze circulations along the entire coast, the central US relies on transport of moisture from the Gulf of Mexico via the low-level jet, and to a lesser extent by moisture transport from the subtropical Pacific during the summer North American monsoon.

2.5 Monsoons and Monsoon Onset Timing Identification

At the heart of large-scale circulations in the tropics are monsoons. The monsoon is a circulation system with seasonal reversal that is driven by the thermal relationships between the ocean and land (Holton, 2004). Monsoon circulations are characterized by a well-defined wet and dry season over the course of a year. Because continental regions have lower heat capacity than the oceans, land regions warm quicker than the oceanic regions for a given amount of solar

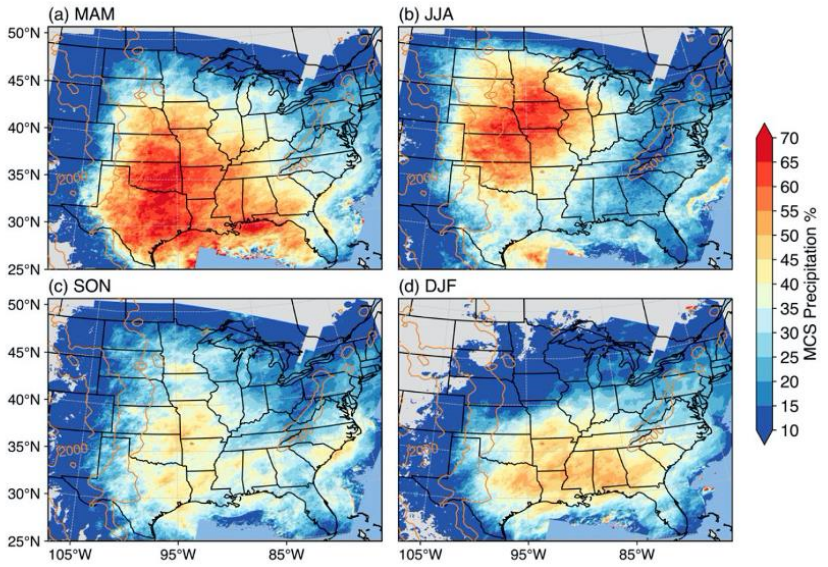


Figure 8: Spatial distribution of MCS rainfall contribution per season from 2004-2016. From Feng et al., 2019.

radiation. Due to the differences in heating and pressure associated with the annual migration of the solar angle, a seasonal continental-scale sea breeze brings moist air inland during the wet season. As a result, warm moist ascent and convective processes are favored during pre-monsoon months, ultimately leading enhanced precipitation during the summer season. One of the driving forces in monsoon onset, including for the South American monsoon, is the location of the Intertropical Convergence Zone or ITCZ (Nieto-Ferreira and Rickenbach, 2011). The ITCZ is normally located over the warmest oceanic waters, which corresponds to each hemisphere's warm season. As such, the ITCZ shifts
southward into the Southern Hemisphere at the conclusion of the northern hemisphere warm season, characterized by weakening of the southern hemisphere cold tongue. The ITCZ itself is a region of low pressure and convergence, making it a prime lifting mechanism. When shifted southward into the southern hemisphere, warm air can rise and cool, resulting in cloud formation and enhanced precipitation during the rainy season.

Several studies have looked at various components related to the monsoon circulations to examine regional onset timing. Wang et al. (2009) objectively defined onset of the Indian summer monsoon by observing the sustained 850-mb zonal wind over a portion of the South Arabian Sea. Their findings suggest that the establishment of westerly flow over the South Arabian Sea is a good indicator of rainy season onset in the Indian peninsula, as 90% of analyzed years have shown agreement between their subjective indicator and onset date established by the Indian Meteorological Department. Kousky (1988) used threshold values of low outgoing longwave radiation (OLR), a proxy for deep convection, as a determining factor in onset date. Marengo et al. (2001) follows this same general framework for the South American monsoon

onset date, except the threshold value used is related to precipitation based on a rainfall climatology. The threshold of onset used in their study was the seasonal average precipitation value of 4 mm day^{-1}. Onset was determined in the following manner. First, the onset pentad must have an average precipitation value greater than the onset value of 4 mm day^{-1}. Next, 6 of the 8 preceding pentads must have an average precipitation rate less than 3.5 mm day^{-1}. Finally, 6 of the 8 succeeding pentads much have an average precipitation rate of 4.5 mm day^{-1}., Nieto-Ferreira and Rickenbach (2011) used a similar onset criterion, using a rainfall threshold to define onset for an array of 5°x5° boxes across the domain of the South American Monsoon System. The rainfall threshold used was 4.6 mm day^{-1}, which is the mean rainfall between 1979 and 2007. The same three general requirements from Marengo et al. (2001) had to be met for onset pentad determination, with the exception that the updated rainfall threshold value was used: (1) The onset pentad must have an average precipitation value greater than 4.6 mm day^{-1}, (2) 6 of the 8 preceding pentads must have an average precipitation rate less than 4.1 mm day^{-1}, and (3) 6 of the 8 succeeding pentads much have an average precipitation rate of 5.1 mm day^{-1}. An exception was also introduced in this study if the succeeding pentad requirement was not met to account for oscillations between active and break phases in the precipitation pattern. During this oscillating phase, if the peak value of 3 out of 8 future pentads doubled (9.2 mm day^{-1}) the threshold value, the pentad of interest would still be determined as the onset pentad (Nieto-Ferreira and Rickenbach, 2011). The implementation of grids in the Marengo et al. (2001) and Nieto-Ferreira and Rickenbach (2011) studies are important for exploring regional onset variability due to various controlling factors, such as topographical variations, surface characteristics, and various atmospheric phenomena. The current study will adapt a similar approach to define the onset of the convective season in the SE US.

CHAPTER 3: METHODOLOGY

3.1 Study Area

The study area focuses on two regions of interest, the Central United States (C US) and Southeast United States (SE US). The selection of the SE US encompasses the same study area from Rickenbach et al. (2020), where the eastern domain extends from Florida to Virginia while the western domain extends from the Gulf Coast to the Ohio Valley. The C US extends from the Mississippi River west to the Front Range of the Rocky Mountains, and north to the Canadian border. While the focus of analysis will be the C US and SE US, this domain enables onset analysis for the Northeast United States as well.

3.2 Data

The Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation (QPE) system produces a state-of-the-art multi-radar and multi-sensor precipitation dataset (Zhang et al., 2016). The main purpose of the MRMS is to create a high spatial and temporal resolution 3D radar mosaic for the CONUS and Southern Canada for both research and operational applications (Zhang et al., 2016). The radar base data is integrated with a multitude of other data, such as atmospheric environmental data, satellite data, rain gauge observations and lightning observations to create various severe weather and QPE products (Zhang et al., 2016). The MRMS system was created from components of two similar systems, the Warning Decision Support System–Integrated Information and National Mosaic and Multi-Sensor QPE (NMQ) systems (Zhang et al., 2011) (Zhang et al., 2016).

Encompassing the CONUS and Southern Canada, the MRMS domain ranges from 20°-55° N latitude and 130°-60° W longitude. The MRMS system primarily relies on a combination of 146 Weather Surveillance Radar 88 Doppler (WSR-88D) and 30 Environment Canada radars, in

addition to 7000 adjusted hourly rain gauge measurements and High-Resolution Rapid Refresh model data, for its precipitation measurements (Figure 9) (Zhang et al., 2016). An advantage of the MRMS QPE dataset is the aforementioned temporal and spatial resolution, which currently stands at 5 minutes temporally and at 1 km North-South x 0.6-1 km East-West spatially (Zhang et al., 2011, Zhang et al., 2016). The radar reflectivity data is quality controlled by using a correlation coefficient filter. Correlation coefficient is a measure of the consistency of various radar targets regarding their shape and size. High correlation coefficient values are found in objects with greater symmetry, like rain and snow. Conversely, lower values of correlation coefficient are found in more non-symmetrical shapes, like animals and buildings. Correlation coefficient radar products are useful because they help identify whether radar targets are meteorological or non-meteorological phenomena. As such, the MRMS correlation coefficient filter separates data into hydrometeor (precipitation) and nonhydrometeor (biological) classifications and removes artifacts

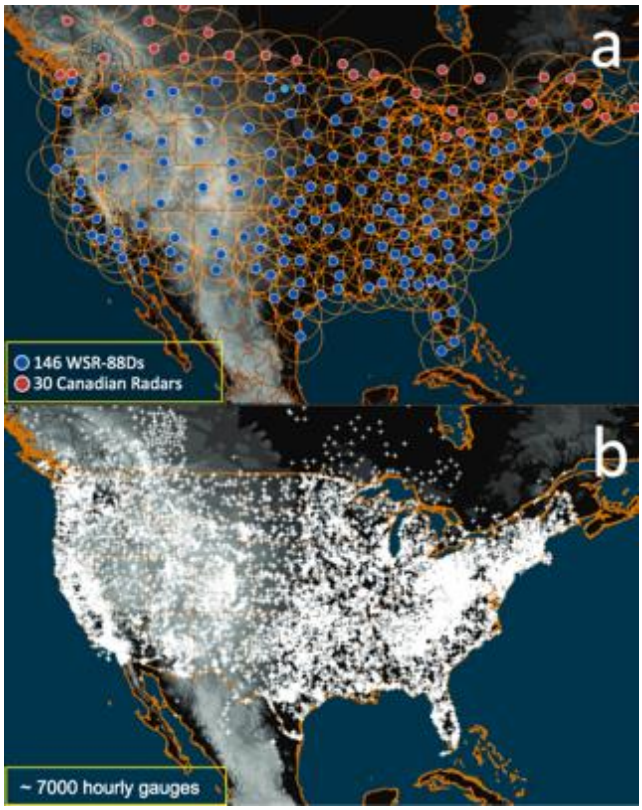


Figure 9: MRMS domain, radar coverage (a), and rain gauge sites (b). From Zhang et al., 2016.

deemed non-hydrometeorological. Three-dimensional reflectivity structure observations are used to retain features like hail, that tend to have low correlation coefficient values due to asymmetries in shape and size and remove biological clutter that have high values for correlation coefficient (Zhang et al. 2016). More than 99% of echoes from nonhydrometeorological phenomena are typically successfully removed using this correlation coefficient quality control method (Zhang et al., 2016).

The MRMS precipitation data for this project is a gauge-adjusted hourly instantaneous rain rate QPE product set temporally between March and July for the years of 2003 through 2011. This hourly instantaneous rain rate QPE product, called PRECIPRATE.HSR, benefits from a five-minute sampling frequency, allowing for more precise quantitative estimation. Since the QPE data being used is instantaneous, the precipitation feature size threshold from Rickenbach et al. (2020) can be implemented, where IPF and MPF feature separation is based around a value of 100 km horizontal dimension. More specifically, precipitation features with horizontal dimension less than 100 km are classified as IPF, while features greater than or equal to 100 km are classified as MPF. Rickenbach et al. (2015) provides an example for the SE US of a group of precipitation features and their corresponding classification as MPF or IPF when the separation threshold length of 100 km is implemented (Figure 10).

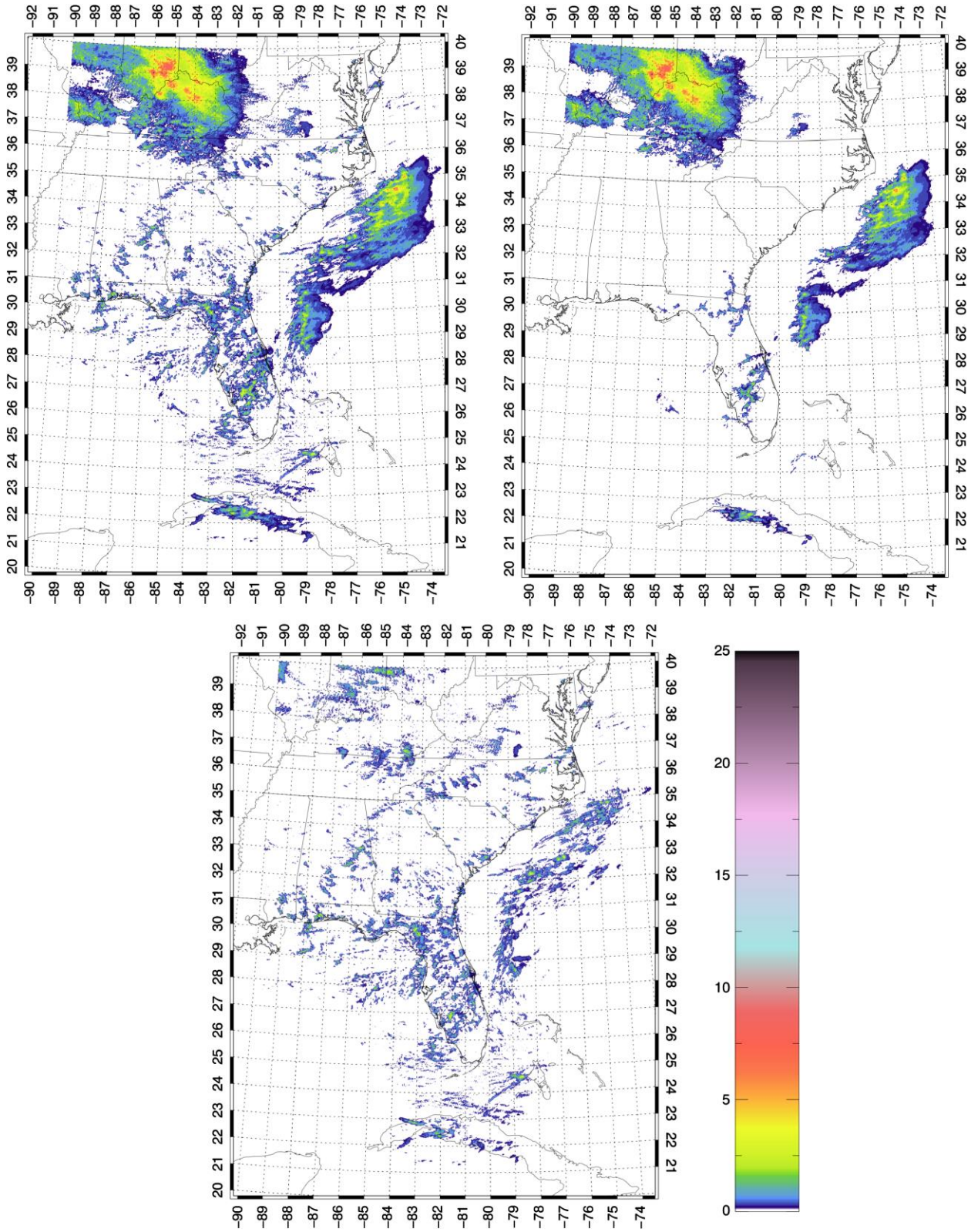


Figure 10: Example of precipitation feature separation into MPF and IPF classifications. The top left represents total precipitation. The top right shows the MPF precipitation. The bottom plot shows IPF precipitation. From Rickenbach et al., 2015.

In order to objectively classify precipitation features as IPF or MPF, each instantaneous hourly QPE data file is processed using an Interactive Data Language (IDL) program that automates the separation process. This separation algorithm will be run on an NCAR Casper Supercomputer. Using Casper, the separation algorithm is submitted as a batch job that takes 72 hours to complete for both higher and lower resolution box arrays, corresponding to how long it takes to process one month of MRMS precipitation data. Using the precipitation data from the top of the hour in each file, the program first identifies which regions of precipitation are rain features based on rain rate, where 0.5 mm hr^{-1} is the defining threshold for light rain (Rickenbach et al., 2015). Using contour border locations of each feature, the maximum length and area of each precipitation feature is measured. The array of border locations is crucial since the maximum lengths of each feature on the array will be used for IPF and MPF classification and eventual summation of precipitation. Measured one feature at a time, features whose maximum horizontal length in one direction is greater than 100 km (converted from degrees of latitude and longitude) are defined as MPF and all others are classified as IPF. Once all features are identified as IPF or MPF, precipitation at each pixel is summed, yielding daily IPF, MPF and total precipitation.

The daily precipitation is used to produce two different types of maps. The first map produced is daily values of IPF, MPF and total precipitation. Figures 11, 12, and 13 show examples of the daily IPF, MPF and daily rain rate plots for 1 April 2011. The second type of map produced using daily precipitation are five-day pentad averages of IPF, MPF and total precipitation. Figures 14, 15, and 16 show examples of the five-day pentad average precipitation for IPF, MPF and total rain for 1-5 April 2011. In both the daily rain rate and pentad average plots, the MPF plots closely resemble the total plots. The larger-scale precipitation features

produce greater quantities of daily precipitation when compared to the smaller, more localized isolated features which ultimately produce less precipitation. As such, MPF rain tends to account

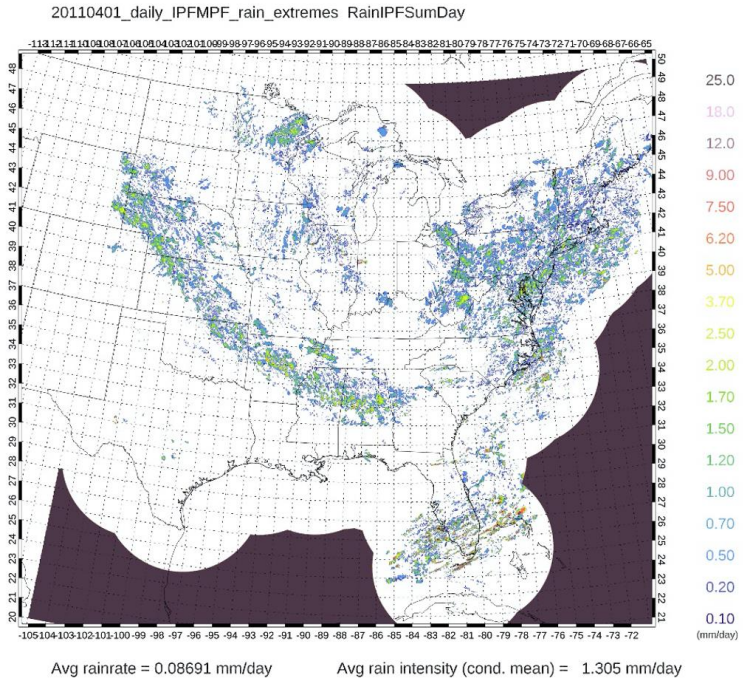


Figure 11: IPF daily rain rate for 1 April 2011.

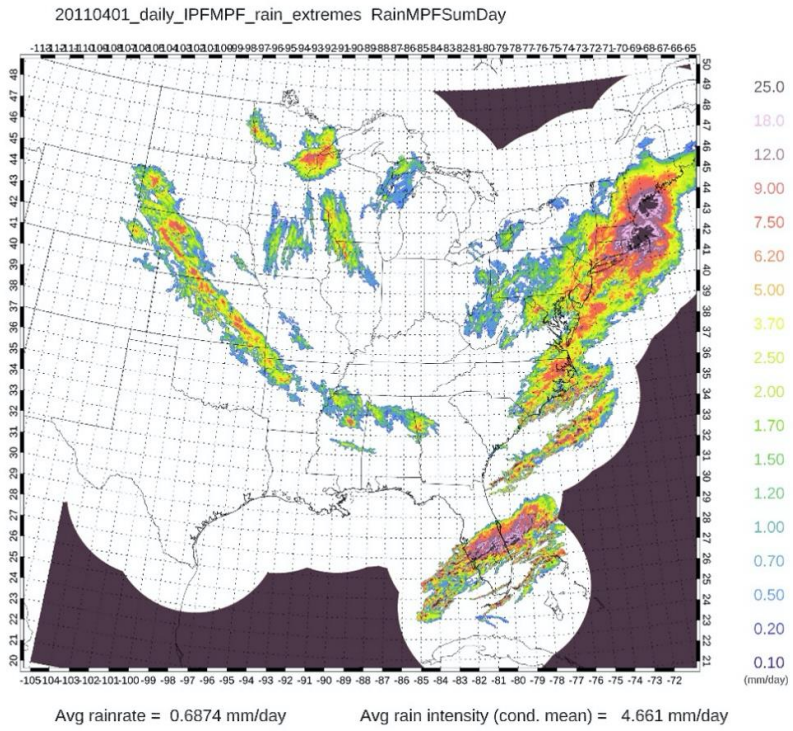


Figure 12: MPF daily rain rate for 1 April 2011.

for a greater portion of total daily rain than that of IPF rain.

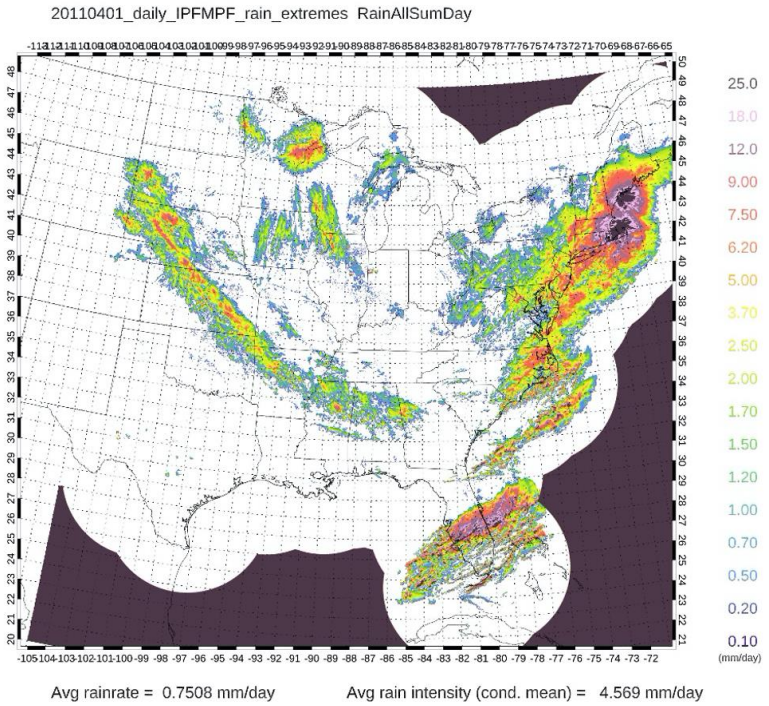


Figure 13: Total daily rain rate for 1 April 2011.

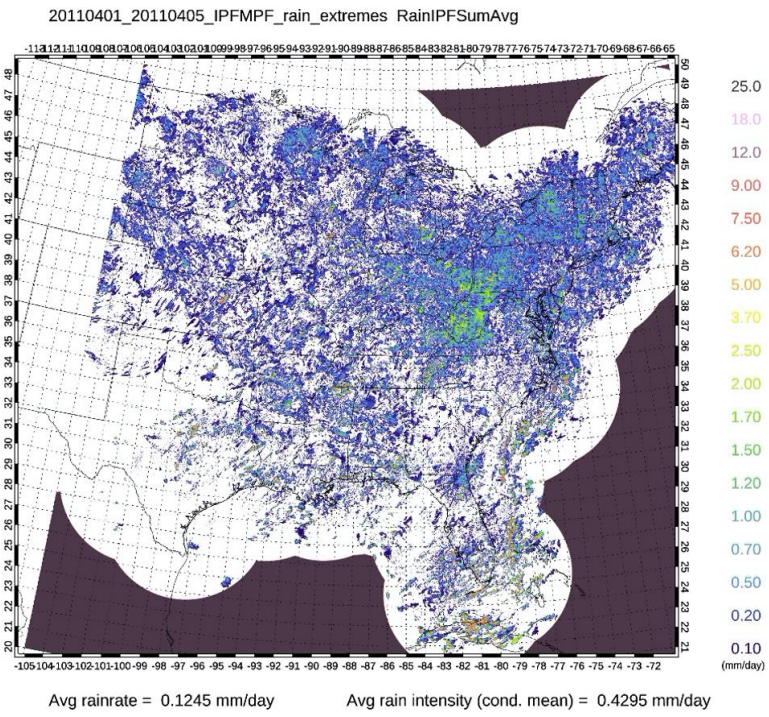


Figure 14: 1-5 April 2011 Pentad average of IPF rain.

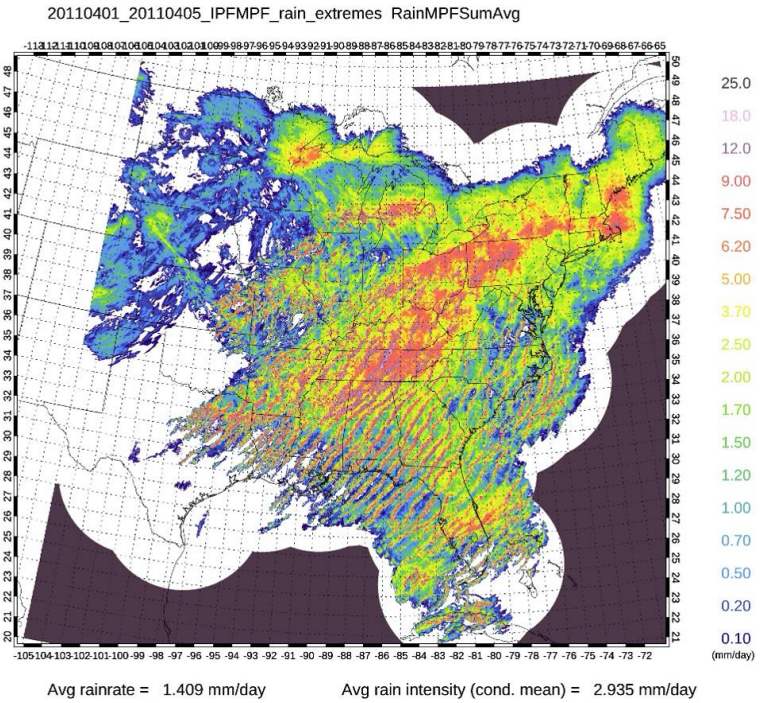


Figure 15: 1-5 April 2011 Pentad average of MPF rain.

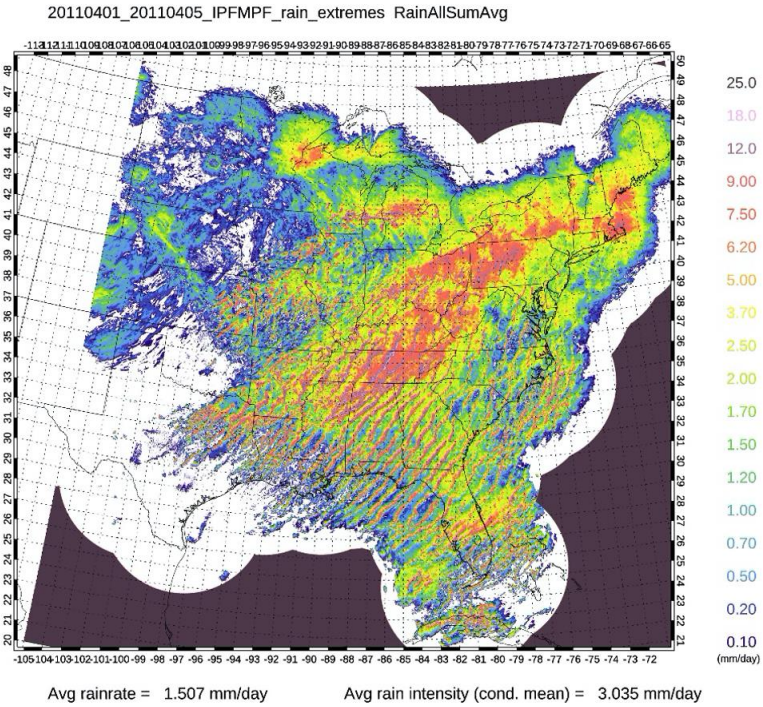


Figure 16: 1-5 April 2011 Pentad average of Total rain.

3.3 Data Analysis

3.3.1 General Overview of Analysis Methods

A method for identifying the springtime onset of IPF precipitation (a proxy for the convective season onset) for the central US and SE US will be performed using the MRMS QPE data. Generally, the methodological approach follows that of Marengo et al. (2001) and Nieto-Ferreira and Rickenbach (2011) who analyzed onset evolution of the South American monsoon based on precipitation thresholds from climatological mean rain rates, for different geographic regions. Rickenbach et al. (2020) adapted their approach, analyzing the geographic and temporal characteristics of the onset of the summer convective regime in the SE US using an array of 2° x 2° boxes covering the domain. The present analysis uses the Rickenbach et al. (2020) methodology and tests the sensitivity of the onset results on the size of the box. The MRMS QPE data will be analyzed each year for the months March – July on hourly, daily, and pentad (five-day) temporal scales using both 2° x 2° (Figure 17) and 4° x 4° (Figure 18) boxes within the domain using IDL programs. With the 2° x 2° box array, analysis can be undertaken at higher resolution, revealing onset characteristics at more localized scales. In using the 4° x 4°, the broader resolution accounts for onset characteristics over a larger spatial area. This is of importance when looking at how onset characteristics change as a result of various geographic features and terrain, like that of oceans and mountainous regions. For each of these studies, onset date is determined using a threshold rain rate value, which is typically a mean precipitation value, and the ability to remain above the threshold value for a specific period after. For this project, the threshold value corresponds to that used in Rickenbach et al. (2020) to allow for direct comparison of the two studies. The implementation of boxes in the analysis allows for regional comparison and differentiation of onset on both temporal and spatial scales. Further, the use of two different box

sizes across the domain will serve as a test of sensitivity related to the temporal and spatial characteristics of convective onset.

3.3.2 Overview of Analysis Methods for the MRMS QPE Data

There are several steps in the MRMS data analysis process to successfully determine the temporal and spatial characteristics of convective season onset for this project. The instantaneous MRMS will first be processed using an IDL program to create output arrays of daily IPF, MPF and total precipitation using the separation algorithm described above. Using an IDL program, arrays of daily average IPF, MPF, and total rain are plotted on the central-eastern US domain as shown in Figures 11, 12 and 13. Next, these daily arrays are used to construct pentad averages of IPF, MPF and total rain across the domain as shown in Figures 14, 15 and 16. The pentad values of IPF rain at each box will be used to create time series of IPF precipitation for March-July of each year across the domain, in addition to time series showing the seven-year average of IPF rain.

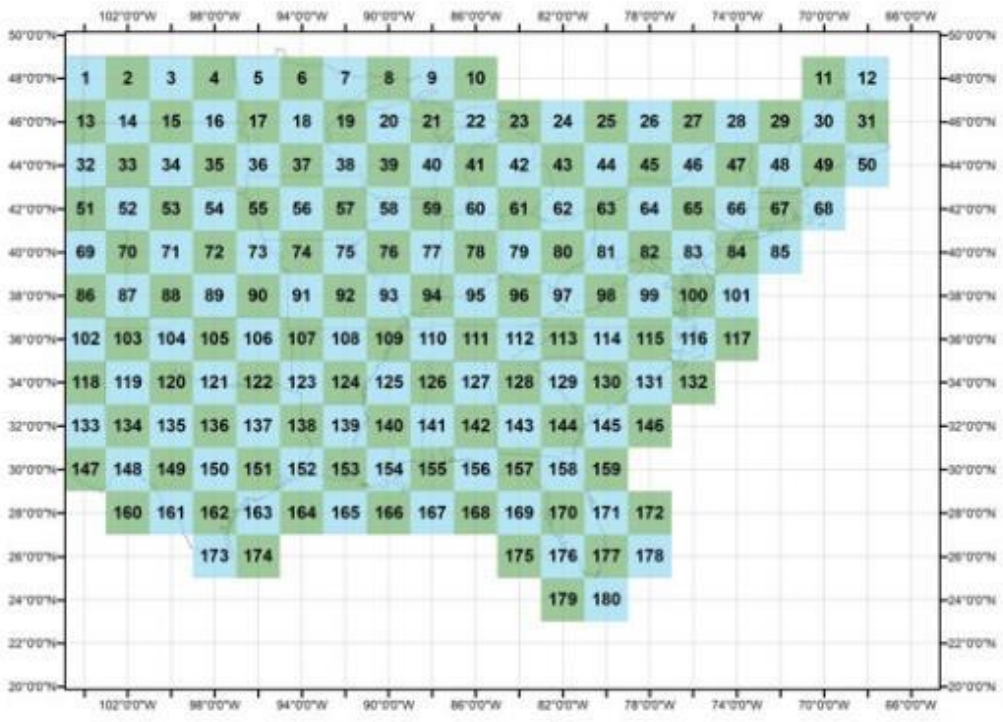


Figure 17: The array of the 2° x 2° boxes across the project domain.

The onset analysis methodology of Rickenbach et al. (2020) will be applied to the IPF time series for both the 2° x 2° and 4° x 4° boxes, though the 2° x 2° box analysis will used primarily as a means of "zooming in" on specific 4° x 4° boxes of interest. Two criteria are required for a pentad to be considered the onset pentad. For each box, the first pentad that meets both of the following criteria during the search is identified as the pentad of onset. The criteria for onset in each box are as follows: (1) The IPF rain value must be greater than or equal to the threshold value, and (2) five of the next eight pentads must have an IPF rain rate greater than or equal to the threshold value. This onset analysis algorithm will be implemented within a Python programming environment. By writing code designed for a Python environment, onset date can be objectively determined for each 2° x 2°and 4° x 4° box in the domain. Pandas, a tool created for a Python programming environment, allows Python to manipulate and analyze numeric tables and time series. Using the Pandas' Python package, the onset analysis algorithm uses Microsoft

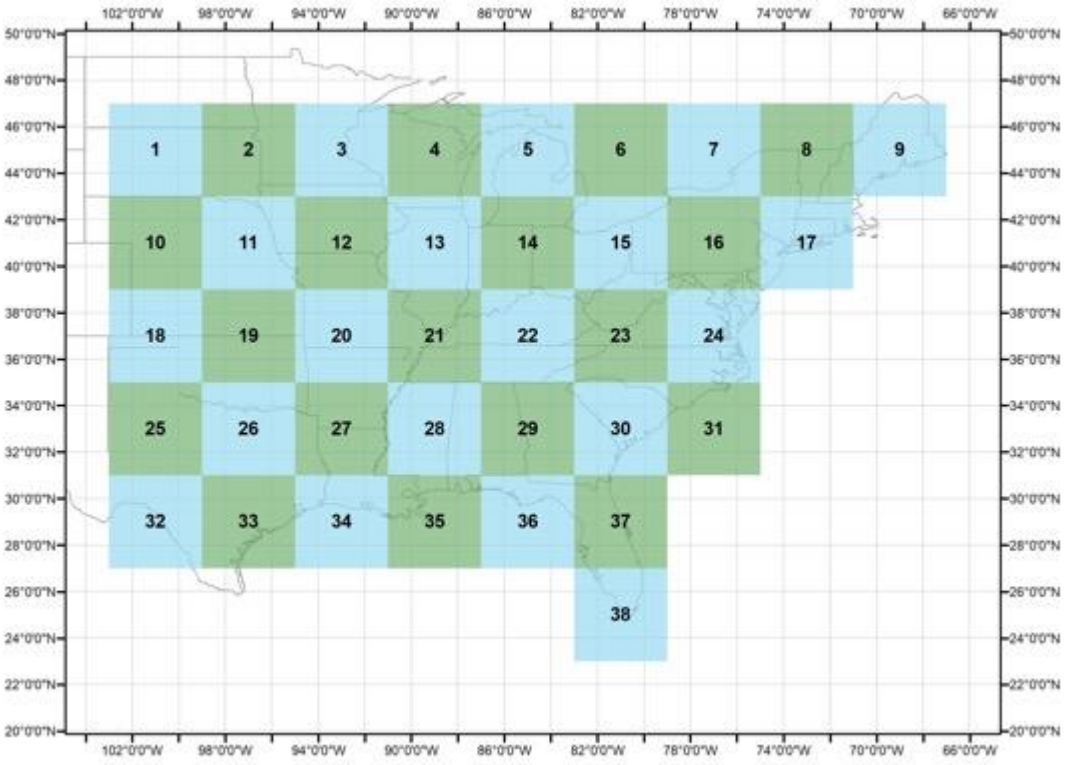


Figure 18: The array of the 4° x 4° boxes across the project domain.

Excel tables containing pentad-averaged rainfall rates for each box as its input and interprets the values with respect to the defined threshold described above. The Python program later writes a new Excel spreadsheet containing "TRUE" and "FALSE" Boolean operators as its output. "TRUE" values correspond to pentads where the rain rate meets or exceeds the defined threshold value. Conversely, "FALSE" values indicate pentads where the rain rate fails to meet or exceed the threshold. Values named "FALSE" are removed from the spreadsheet by searching for and replacing all "FALSE" values with three underscores. This inherently leaves only "TRUE" values within the spreadsheet. These values are manually examined to determine the pentads in which the onset criteria is met for each box. The same criteria described above is used to determine onset in this manner, where both a single pentad's rain value and five of the next eight subsequent pentads rain rates must exceed the threshold. Since "TRUE" values are defined as pentads having met or exceeded the threshold, onset in this manner is now met when both a single pentad and five of the eight next pentads are defined as "TRUE". The year-to-year consistency of the onset pentad for each box will be measured by looking at the standard deviation of onset at each box over the ten-year period. For a given box, lower standard deviation values correspond to greater consistency in the pentad of onset from year-to-year. Conversely, greater standard deviation values correspond to greater variability in the pentad of onset from year-to-year. The methodological flow of the data processing and analysis processes related to onset date determination are shown in Figure 19.

Using the processing stream described above, the onset pentad for each year and for all years combined in each of the 2° x 2° boxes and the 4° x 4° boxes was determined. From these findings, maps were produced showing pentad and monthly IPF precipitation, along with maps

showing the onset pentads for both box arrays. Presentation and interpretation of these maps will be done in the Results and Discussion chapter.

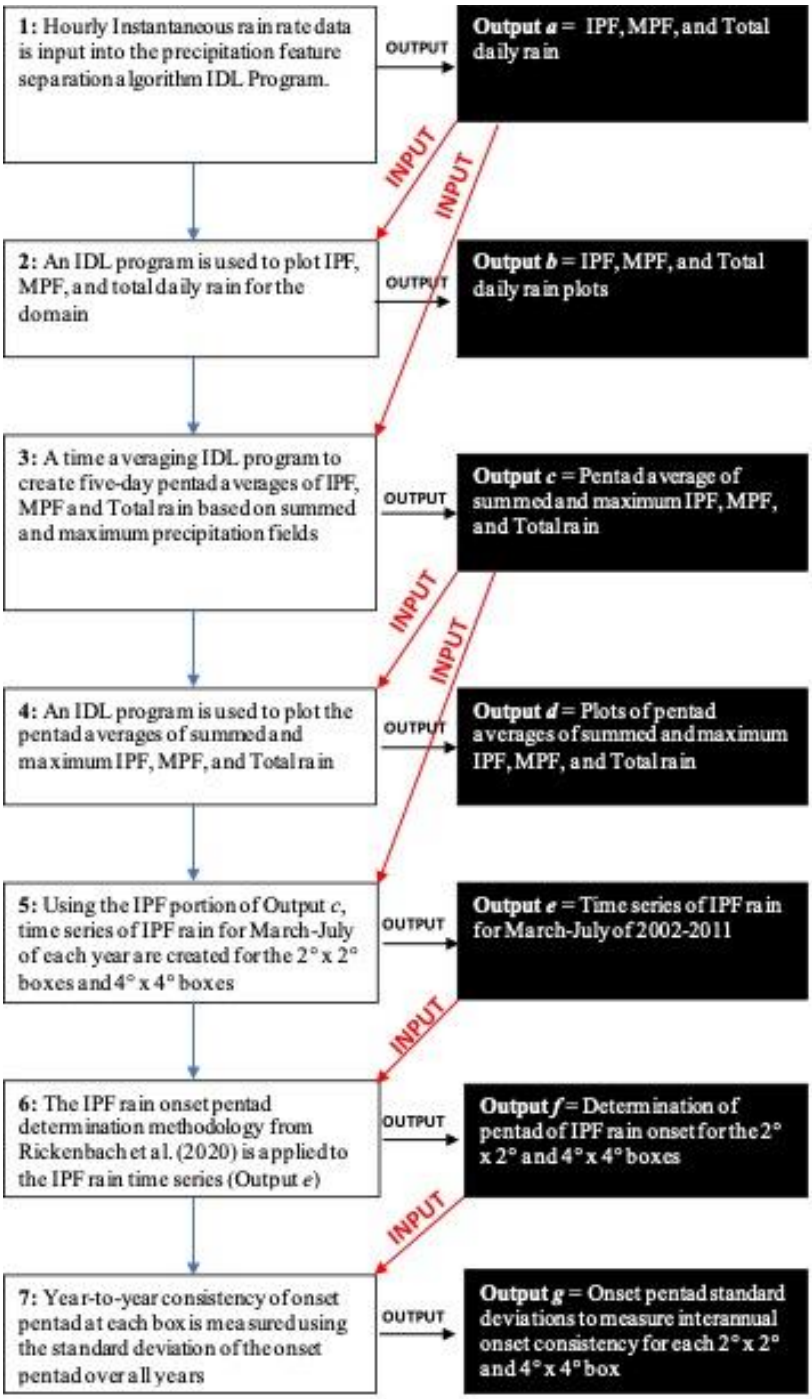


Figure 19: Flowchart of data processing and analysis in determining the pentads of isolated convective season onset.

CHAPTER 4: RESULTS AND DISCUSSION

4.1 Data Quality

Before discussing the onset results, important data quality issues for two of the years are presented first. The hourly instantaneous MRMS QPE dataset used for this analysis was created and acquired from the National Oceanic and Atmospheric Administration (NOAA) National Centers for Environmental Information (NCEI) in Asheville, NC. As is often the case with large datasets, missing or corrupted is a possibility in large-scale data collection and transfer. This was the case in some cases for the analysis of this study (Figure 20). In particular, there are noticeable portions of data not available for the years of 2008 and 2009.

For 2008, the dataset used only contains approximately 61% of data for April and 29% of data for June (Figure 20). As mentioned previously, the data issues for 2008 are obvious when looking at the onset maps, which are shown in section 4.3.2.8 In all but two boxes of the onset domain for 2008 onset is observed to have occurred at either pentad 36 or 37, both corresponding with the end of June in the pentad date chart (Table 1). In the pentads prior to this (pentads 31-35), the five-day average IPF rainfall at many boxes is near or equal to 0 mm day^{-1}, further signifying the absence of this data. If this missing data was included in the dataset, it is possible that onset would occur closer to the seven-year average onset date at pentad 33. Nonetheless, without this data we cannot confidently make this claim.

In 2009, domain-average onset occurs at pentad 29 (May 21-25). Like in 2008, the missing data for 2009 also comes from the months of April and June. It would be easy to make the claim that the domain-average onset date would remain unchanged for 2009, as 100% of May data is included in the dataset. However, when looking at the 2009 onset map shown later in section 4.3.2.9 it is clear that this is not the case. In particular, there are three boxes (boxes 15, 34

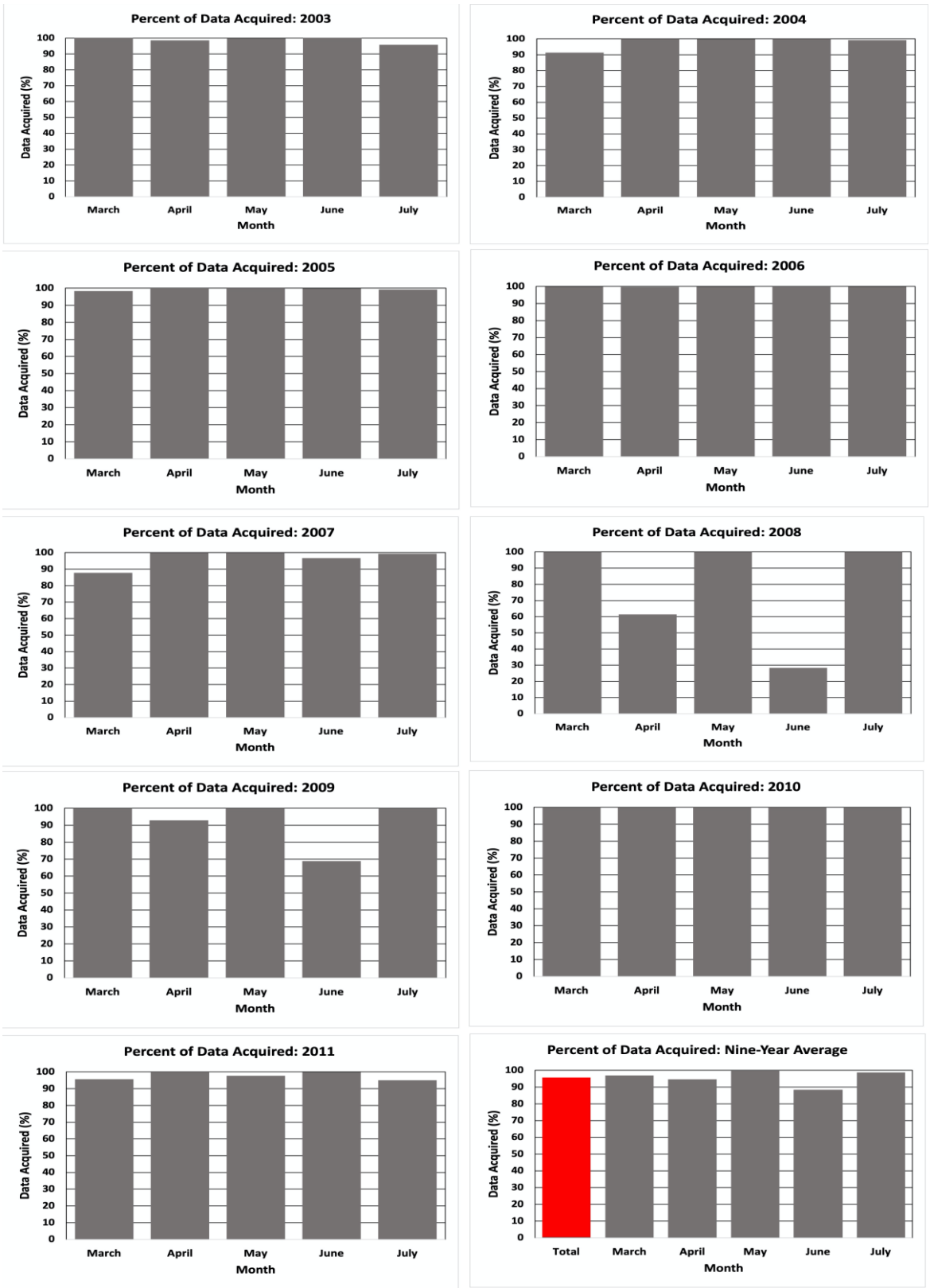


Figure 20: Bar graphs corresponding to the percentage of hourly instantaneous MRMS data available for each month and year of the analysis.

and 35) with an onset date of pentad 36. As with 2008, portions of the early and middle portions of June have a five-day average IPF rain values at or near 0 mm day $^{-1}$. Given these findings and the results in Figure 20, the following discussion of onset results will not include the years 2008 and 2009 when looking directly at average onset. The average onset analysis will only include 2003-2007 and 2010-2011. As such, the years 2008 and 2009 will not be discussed.

Pentad	Date
13	3/2 - 3/6
14	3/7 - 3/11
15	3/12 - 3/16
16	3/17 - 3/21
17	3/22 - 3/26
18	3/27 - 3/31
19	4/1 - 4/5
20	4/6 - 4/10
21	4/11 - 4/15
22	4/16 - 4/20
23	4/21 - 4/25
24	4/26 - 4/30
25	5/1 - 5/5
26	5/6 - 5/10
27	5/11 - 5/15
28	5/16 - 5/20
29	5/21 - 5/25
30	5/26 - 5/30
31	5/31 - 6/4
32	6/5 - 6/9
33	6/10 - 6/14
34	6/15 - 6/19
35	6/20 - 6/24
36	6/25 - 6/29
37	6/30 - 7/4
38	7/5 - 7/9
39	7/10 - 7/14
40	7/15 - 7/19
41	7/20 - 7/24
42	7/25 - 7/29
43	7/30 - 8/3

Table 1: Chart containing pentad numbers and their corresponding traditional calendar dates.

4.2 Local Analysis Region Area Sensitivity

To test the sensitivity of onset timing against areal changes of analysis region, onset results from both the 4° x 4° and 2° x 2° box array domains were used for each year (Figure 21, 52-59) and for the seven-year average (Figure 29). This analysis addresses two questions. The first is how the domain average dates of onset compare between the higher and lower resolution box arrays. The second is to determine if there is any important details evident regarding the geographic pattern of onset in the higher-resolution 2° x 2° array compared to the 4° x 4° array.

Regarding the first question, the timing of the domain-average onset pentad is unchanged in six of the seven years (Figure 21), with a two-pentad difference in 2011. There is also a three-pentad difference in the domain-average onset date between the 4° x 4° and 2° x 2° box arrays for the seven-year average (Figure 29). Given these results, the time of domain average onset does not vary much for this dataset. Analysis of onset sensitivity to the boxes themselves paints a bit of a different picture and is the basis for answering the second question regarding differences between the two different box arrays. In general, the onset date for the 4° x 4° boxes is influenced by the smaller 2° x 2° analysis regions. Due in part to being smaller and more localized in scale, onset date can vary more from box to box on the 2° x 2° than on the 4° x 4° box array. As an example, when looking at 2003, onset at box 31 of the 4° x 4° array occurs at pentad 15. This particular 4° x 4° box encompasses the following 2° x 2° boxes: boxes 116, 117, and 132. Onset at these 2° x 2° boxes differs by as much as 17 pentads, as onset occurs at pentads 27, 15 and 32 respectively. This is only one example of many that exhibits how changing 4° box 38) in 2011. The corresponding 2° x 2° analysis regions (boxes 176, 177, 179 and 180) have onset occurrence at pentads 24, 19, 31 and 28. Conversely, onset for box 38 of the 4° x 4° array occurs at pentad 33. As such, areal changes to the analysis regions also means that IPF

rainfall values vary with these areal changes, also explaining the unequal weighting of the influence that the smaller analysis regions have with respect to the larger regions.

Another interesting point to be made is that the smaller analysis regions tend to show a greater occurrence of isolated onset events. When looking back at the onset analysis of the 4° x 4° box arrays, there was only one year in which isolated onset events were observed (2009). However, when looking at the onset results for the 2° x 2° box arrays, there are isolated occurrences of onset present for all but one year (2004) that are not related to or bordering a larger region of onset occurrence. This phenomenon is also observed in seven-year average onset map for the 2° x 2° box array. Because the larger areal analysis regions rely on multiple smaller analysis areas, many of the 4° x 4° boxes do not have determinable onset. However, one of the four 2° x 2° boxes that is encompassed within this larger analysis region may have determinable onset but can only be observed when looking at the 2° x 2° array. Some examples of this are

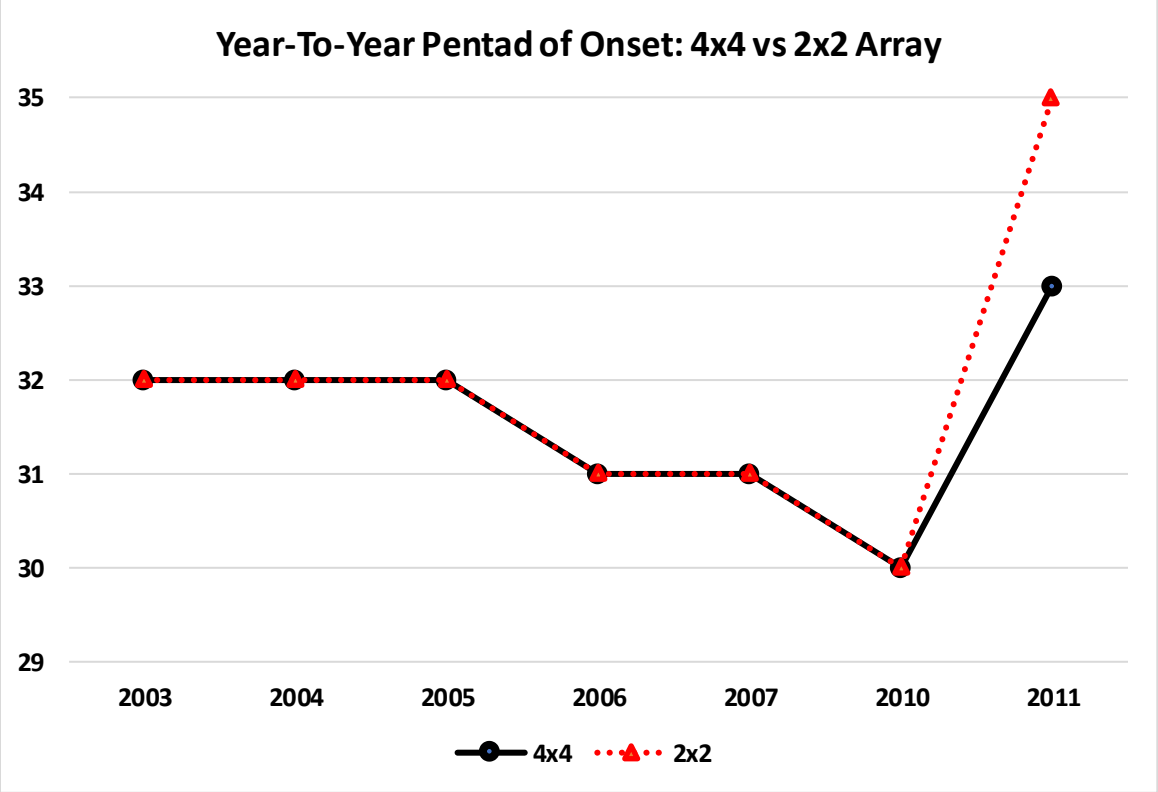


Figure 21: Onset dates for 4x4 and the 2x2 box arrays.

onset occurrence in Minnesota in 2003, Wisconsin, New Mexico and Colorado in 2005, South Dakota in 2007, and Iowa in 2011. In each of these cases, IPF onset is present for one or two

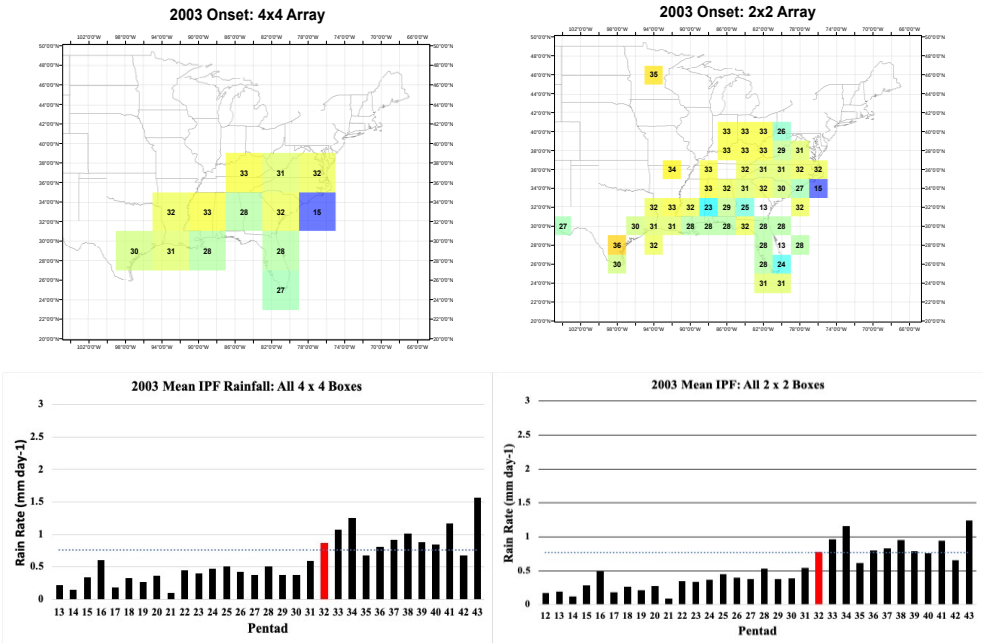


Figure 22: Onset maps and corresponding time series for the 2° x 2° and 4° x 4° box array domains used to test areal sensitivity of onset timing for 2003.

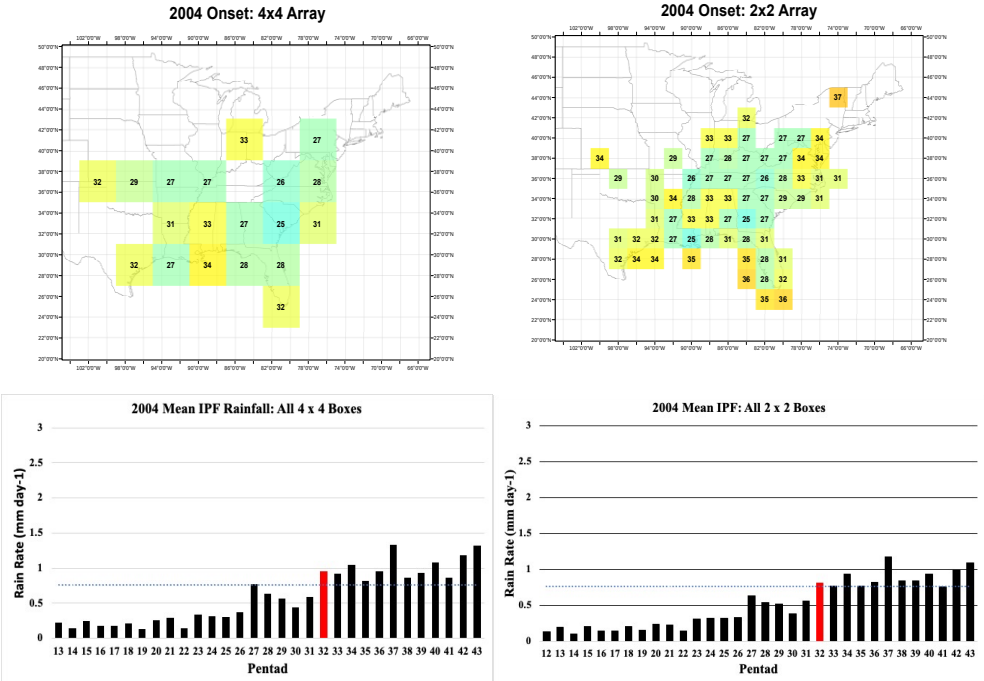


Figure 23: Same as Figure 22, but for 2004.

isolated 2° x 2° boxes but is not picked up in the 4° x 4° onset analysis due to the areal differences in box size. As an example, shown in Figure 30 is a comparison of the 2005 time

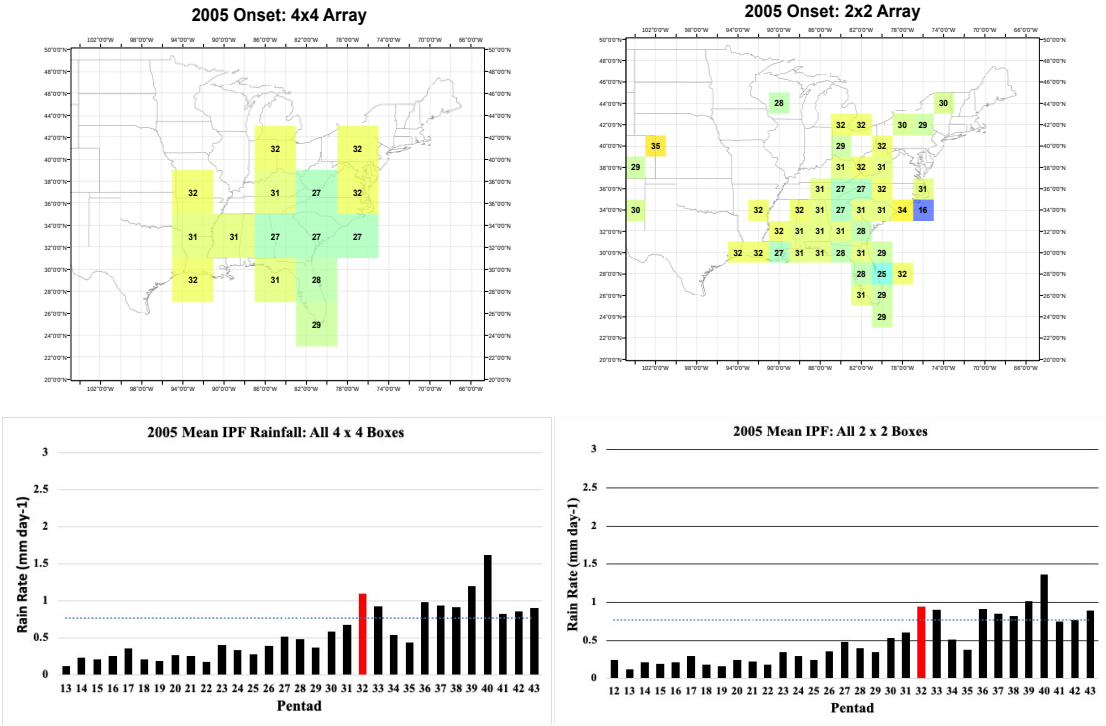


Figure 24: Same as Figure 22, but for 2005.

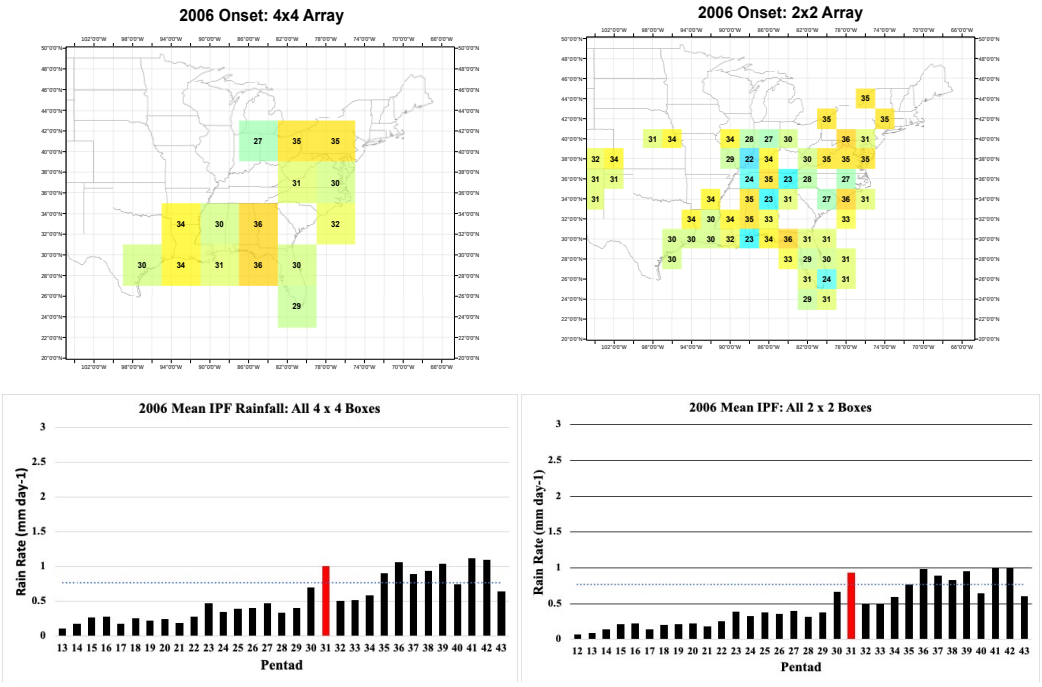


Figure 25: Same as Figure 22, but for 2006.

series for box 10 of the 4° x 4° array and box 70 of the 2° x 2° array, both corresponding to northeast Colorado and western Nebraska. At the larger box 10 of the 4° x 4° array, the onset

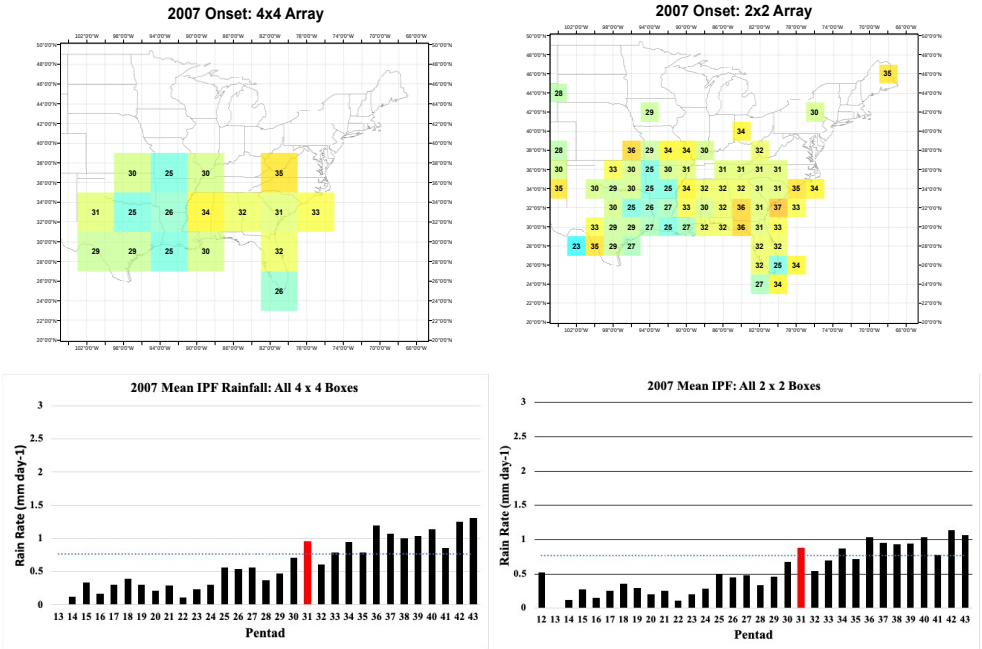


Figure 26: Same as Figure 22, but for 2007.

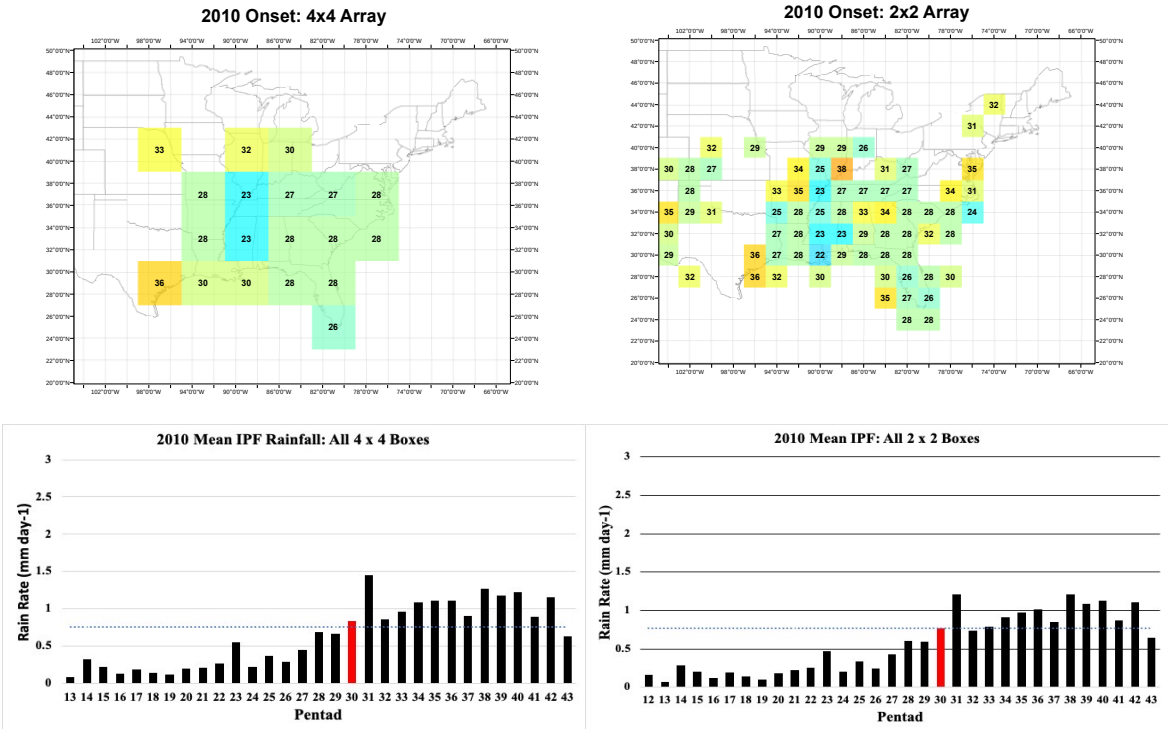


Figure 27: Same as Figure 22, but for 2010.

criteria are not met. Conversely, at the more localized box 70 of the 2° x 2° array, the onset criteria are met at pentad 35. As such, this example serves as evidence that the smaller 2 ° x 2°

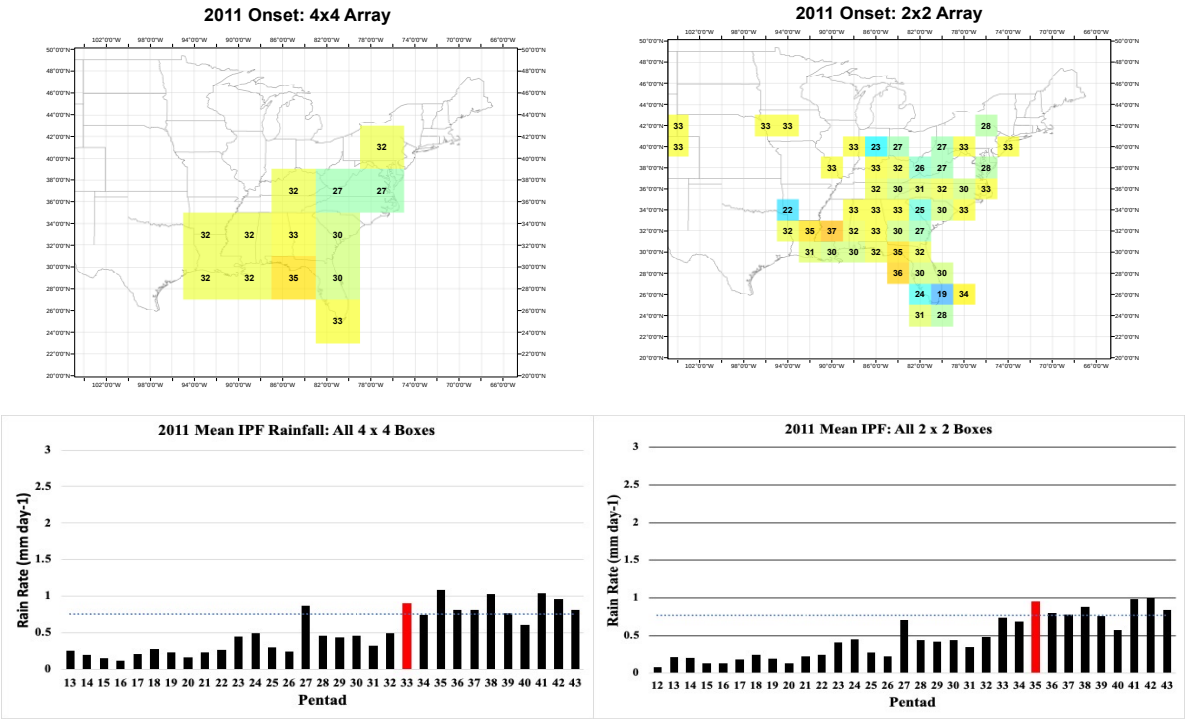


Figure 28: Same as Figure 22, but for 2011.

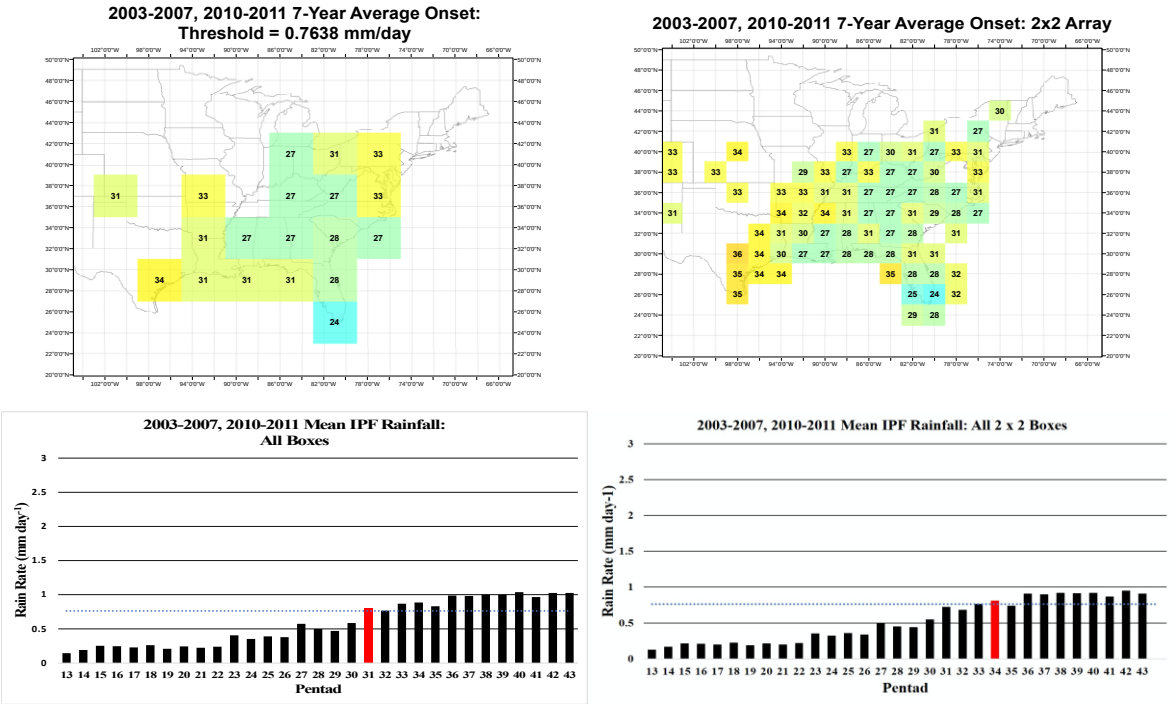


Figure 29: Same as Figure 22, but for the seven-year average.

boxes are subject to more localized or random onset occurrences, while the larger 4° x 4° boxes often rely on more widespread onset occurrence within multiple of the smaller boxes. However, since the trends and patterns observed in both the 4° x 4° and 2° x 2° boxes are generally the same, with the exception of some localized differences in the 2° x 2° array, using the 4° x 4° boxes for the onset analysis is sufficient.

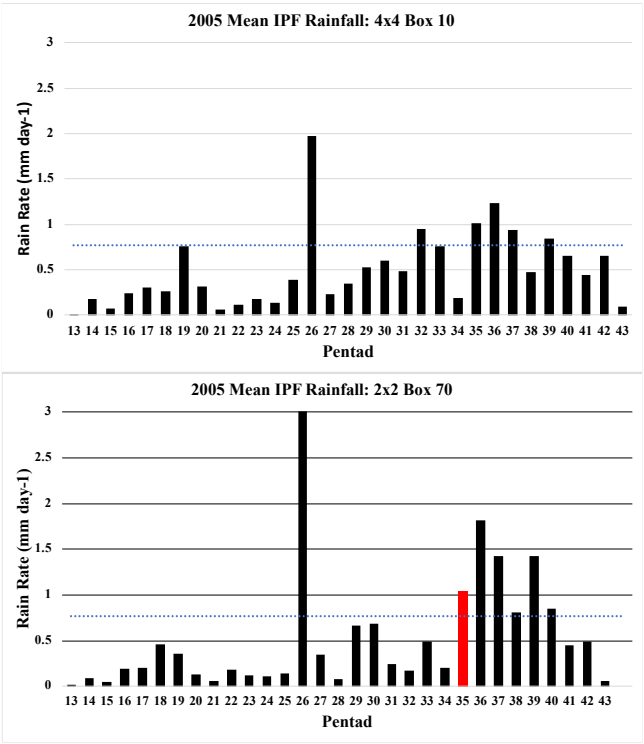


Figure 30: Comparison of onset dates at 4° x 4° box 10 and 2° x 2° box 70 in 2005.

4.3 Results of IPF Onset Determination

This section presents the primary results of the current study. It begins with an overview of the average IPF onset for the seven-year period, followed by year-to-year characteristics of onset. Next will be a discussion of the spatial and temporal behavior of onset, based on both the seven-year average and year-to-year characteristics of onset. Following this will be a discussion of onset sensitivity with respect to two factors: (1) local analysis region area and (2) the rainfall threshold used to determine onset. Finally, this section will conclude with a discussion related to data quality and the role it played in onset determination results.

4.3.1 2003-2011 Nine Year Average Onset

For the period encompassing 2003 through 2011, onset on average first occurs in south Florida. In this area, onset occurs at the end of April on pentad 24 (April 26-30). Along the SE US, there tends to be a gradual northward and westward progression of onset occurrence (Figure 31). Onset occurs next along much of the Southeast Coastal Plain, Gulf States coastal plain, and northward to the Ohio River Valley in mid-May (pentads 27 and 28). In late May – early June (pentad 31), onset occurs simultaneously along the Gulf coast region. Onset then moves northward from northeastern North Carolina to the mid-Atlantic region, occurring between early-to-mid June (pentads 31-33). The latest onset occurs along the Texas Gulf coast in mid-June (pentad 34). For the entire domain, onset on average occurs in late May – early June (pentad 31) as seen in the seven-year average IPF time series for all 38 boxes combined (Figure 32). For the entire region, Figure 32 shows a steady increase in IPF rain throughout the spring season, rather than a rapid increase for a given pentad.

The northward and westward progression of IPF onset can also observed in monthly average composite rainfall maps of IPF precipitation across the SE US (Figure 33). Beginning with March, the entire domain tends to have very little IPF rain, in the magnitude of 0.1-0.5 mm day^{-1}. In April, the domain still has a relatively small average IPF rainfall presence, though IPF rainfall values across the entire domain are more closely approaching 0.5 mm day^{-1} than in March. Noticeable changes in average IPF precipitation become apparent beginning in the month of May, particularly along the Southeast Coastal Plain region up into the Carolinas. By May, there is a noticeable large increase in IPF rainfall in southern Florida, increasing to upwards of about 3 mm day^{-1}. The increase in IPF precipitation from April to May along the Southeast Coastal Plain and Carolinas is not as pronounced as that of southern Florida, but values in this

region still exceed 1 mm day^{-1} both along the coast and in the Appalachian Mountains. The month of June shows further progression of IPF rainfall across the domain, with noticeable increases now along the Gulf Coast and Southern Coastal Plain regions, in addition to continuing progression along the entire Southeast Coastal Plain up into Virginia. An interesting feature in both the composites for June and July precipitation and the map of average onset is that very little IPF rain is present in the Central Plains region, with the exception of the southern two-thirds of the Texas Gulf Coast. A possible explanation for this is that the springtime increase in moisture and instability from the seasonal establishment of the NASH that results in IPF onset is mainly limited to the SE US (Nieto-Ferreira and Rickenbach, 2021). The seven-year average onset map and monthly average IPF rainfall composite maps both support this idea, suggesting that IPF onset appears to be a phenomenon limited to the SE US. To show this, a map showing the difference in July IPF rainfall and March IPF rainfall was created for the seven-year period

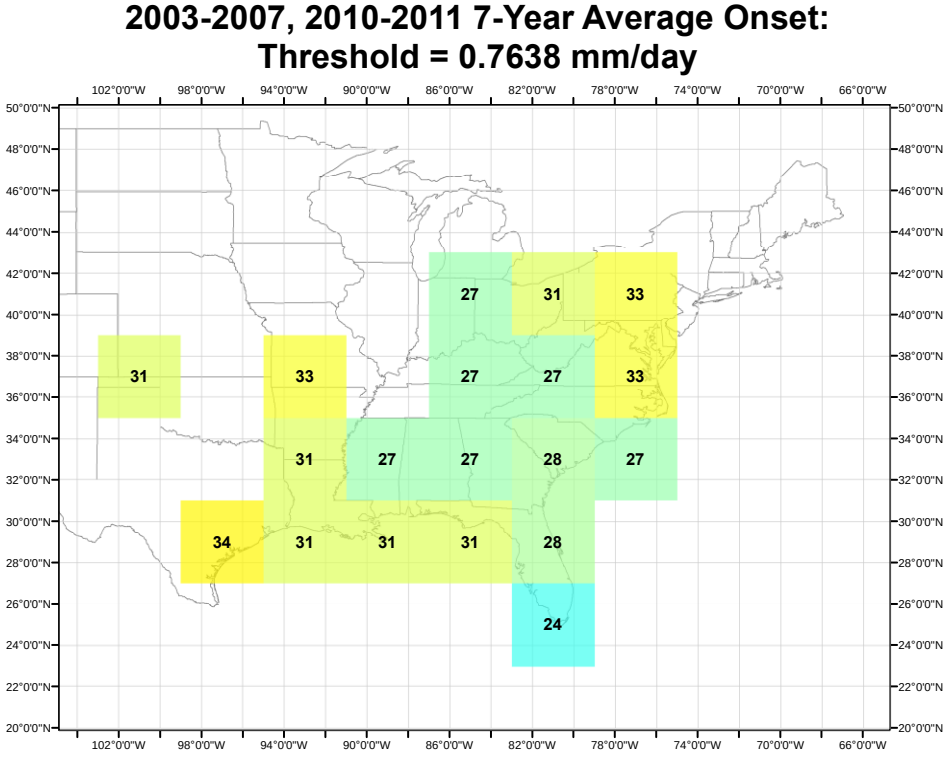


Figure 31: Seven-year average onset across the 4° x 4° domain. Domain-average onset occurs at pentad 33.

(Figure 34). In the SE US where onset occurs (Figure 31), there is a clear increase in IPF precipitation from March to July. The greatest differences are observed across the contiguous SE US, with the July and March IPF precipitation difference typically exceeding 1 mm day^{-1}. For the remaining boxes in the array where onset does not occur, primarily in the Central Plains and the northern portion of the domain, the July minus March difference in IPF precipitation is less than or equal to 0.63 mm day^{-1}, therefore suggesting that the higher rainfall difference (July minus March) is consistent with an increase in moisture availability over the same period as shown by Nieto-Ferriera and Rickenbach (2021).

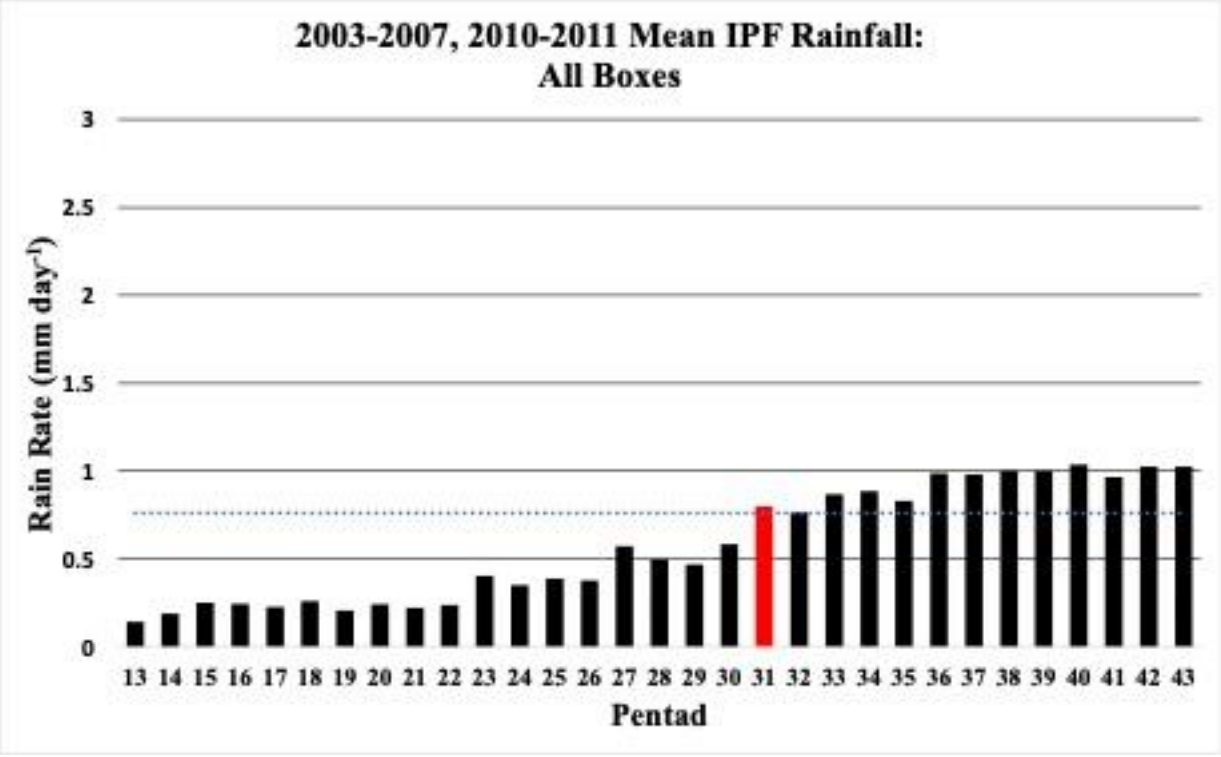


Figure 32: Seven-year average IPF rainfall time series for all 4° x 4° boxes. Domain-average onset occurs at pentad 31.

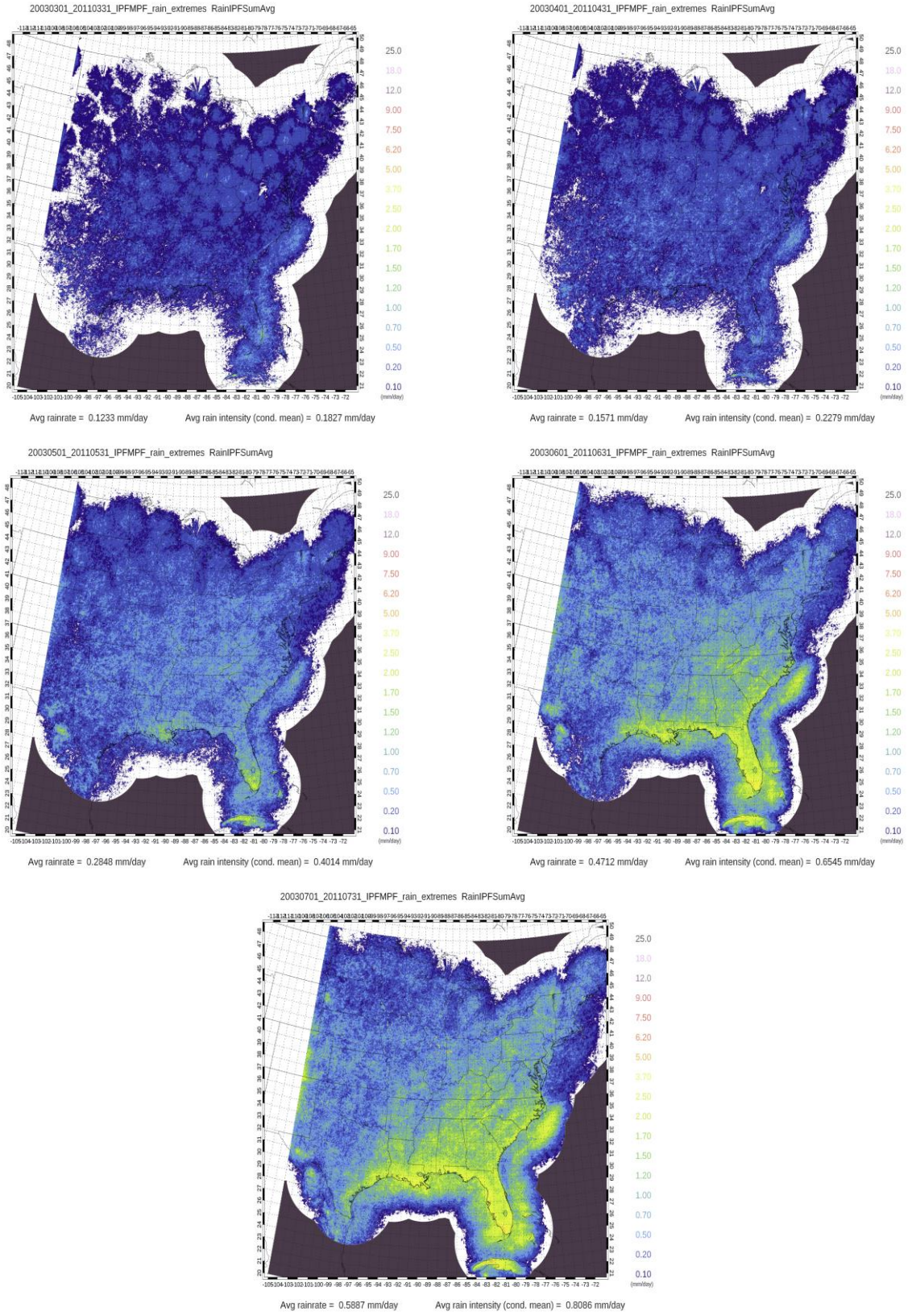


Figure 33: Monthly composite IPF rainfall maps averaged over the seven-year period.

4.3.2 2003-2011 Year-to-Year Onset

The following sections will introduce the findings regarding year-to-year onset date across 4° x 4° box array. The aim here is to examine qualitatively whether there is consistency in the geographic and temporal pattern of onset over the seven-year analysis period. To do this, onset maps for each year are presented below. These maps were created to make it easier to characterize onset across the domain both spatially and temporally for the purpose of comparing year-to-year onset characteristics.

4.3.2.1 2003

Beginning with 2003, onset occurs in about one-third of the boxes, limited to the SE US region (Figure 35). The boxes with onset in 2003 also had onset occurrence when looking at seven-year average onset (Figure 31). In general, similar trends in onset progression are observed for 2003 as were for the seven-year average, where IPF onset (Figure 35) and mean monthly IPF rainfall (Figure 36) gradually progress northward with time. Additionally, onset is limited to

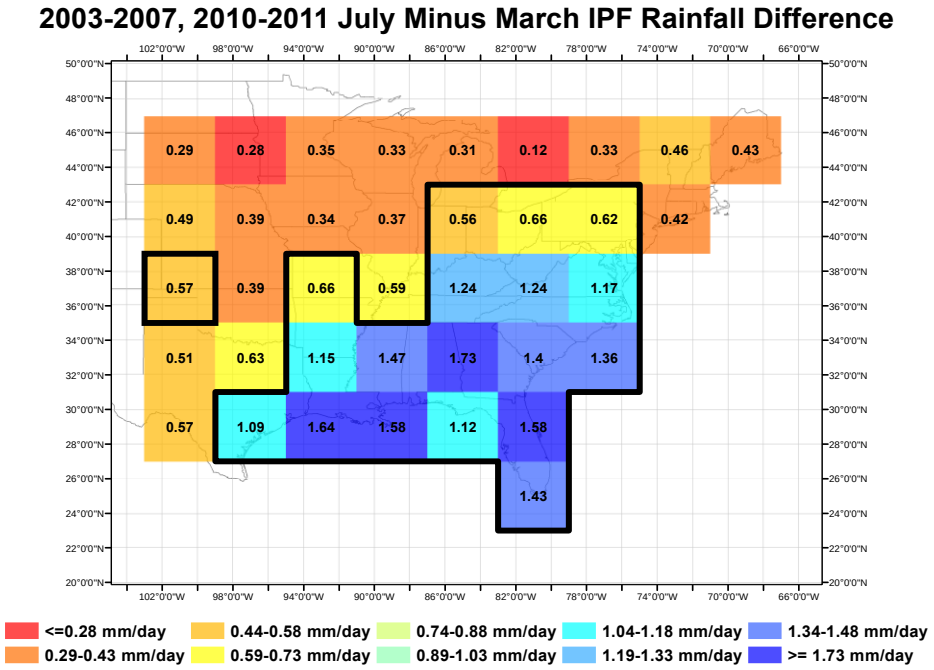


Figure 34: Seven-year average July IPF rainfall subtracted by March IPF rainfall. Outlined in black is the region where onset occurs on average over the seven-year period. Units are in mm day^{-1}.

strictly to the SE US, as is also mostly the case for the seven-year average. While these general trends are consistent with the seven-year average, there are some notable differences in onset for 2003. Onset in the Carolinas generally occurs 1-5 pentads later in 2003 when compared to the seven-year average, with the exception of southeastern North Carolina (box 31) in mid-March (pentad 15). There also exists a five-pentad difference in onset date for the southeast quarter of Texas when comparing 2003 and seven-year average onset. Also, of interest is that the Florida panhandle region (box 36) does not attain the onset criteria. As will be discussed later, 2003 is not the only year where onset is absent in this area. For 2003, the domain-average date of onset is pentad 32, one pentad earlier than the seven-year domain average.

4.3.2.2 2004

While the trends and observations of onset for 2003 were similar to the seven-year average, this is not the case in 2004. In 2004, half of the boxes across the domain have

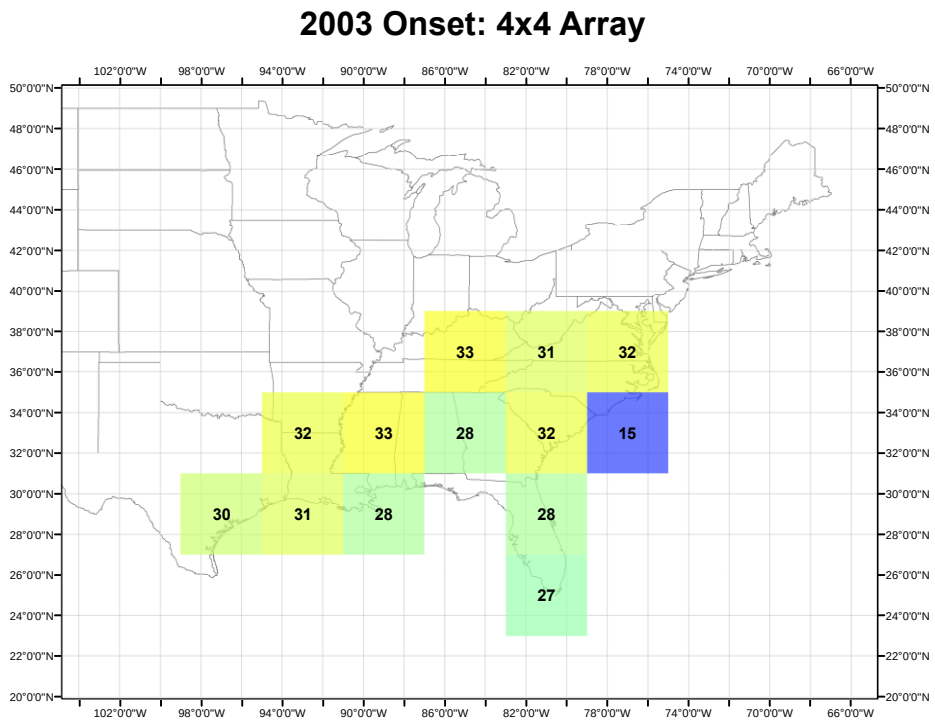


Figure 35: IPF Onset map for the year 2003. Domain-average onset occurs at pentad 32.

determinable onset using the criteria of this study, across the Central Plains, SE US and as far north as New York and Michigan (Figure 37). Of these, in two of the boxes, there is no onset

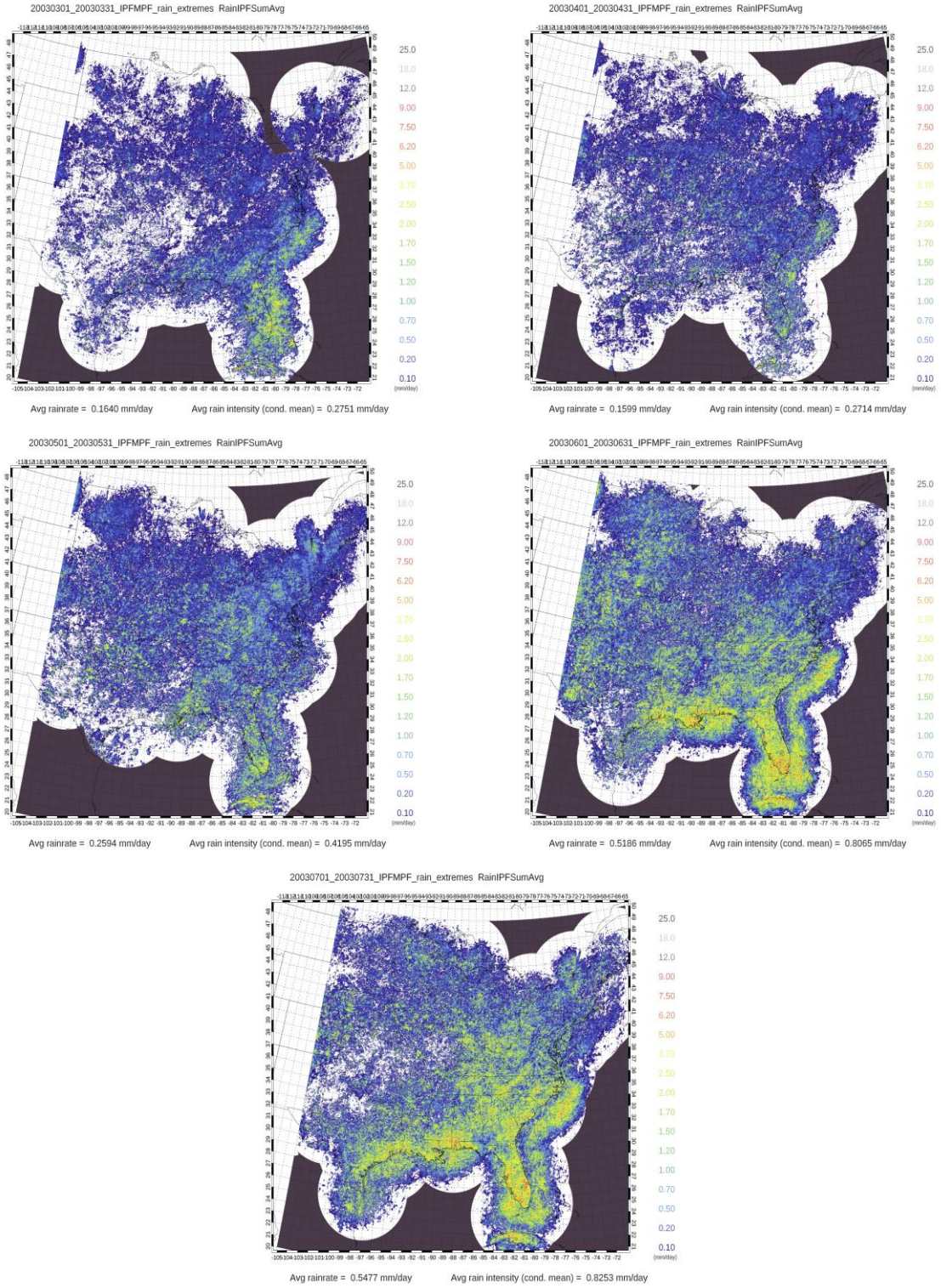


Figure 36: Monthly composite IPF rainfall maps for the year 2003.

defined for the seven-year average at these locations (boxes 19 and 21). Interestingly, onset and monthly IPF precipitation (Figure 38) evolution do not seem to progress northward with time for this year, unlike the seven-year average onset progression. Rather, the earliest occurrences of IPF onset are more sporadically spread out across the domain. The earliest observed date of onset (pentad 25, or early May) is in South Carolina (box 30), and is immediately followed by onset in western North Carolina, Virginia, and south-central West Virginia one pentad later. Unlike for the seven-year mean, this suggests another mechanism besides seasonal southerly moisture advection was active in 2004, perhaps related to a specific weather event centered in the Carolinas. and New York. Florida, where onset is generally first observed, experiences a late start comparative to the Carolinas, with onset occurring in mid-May to early June. A noticeable

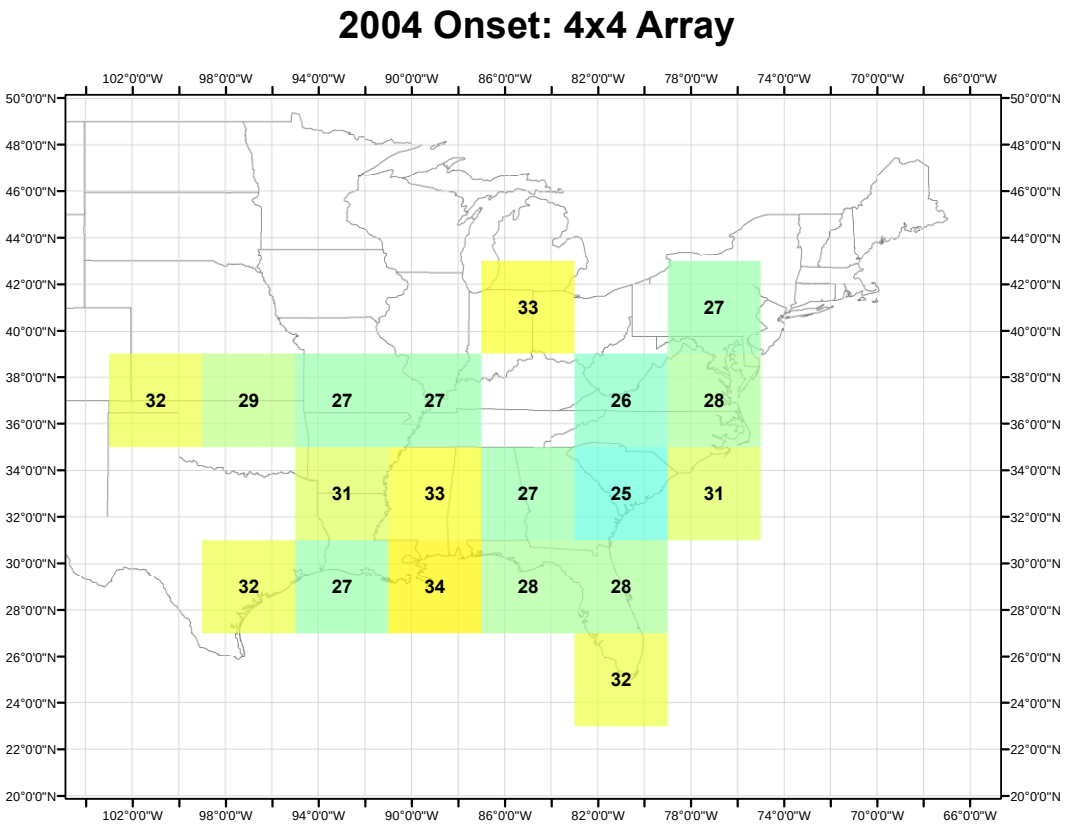


Figure 37: IPF Onset map for the year 2004. Domain-average onset occurs at pentad 32.

feature present in 2004 is the expansion of IPF onset into the central plains, with mid-May to early-June onset occurrence across north-central Oklahoma, south-central Kansas, and the Texas

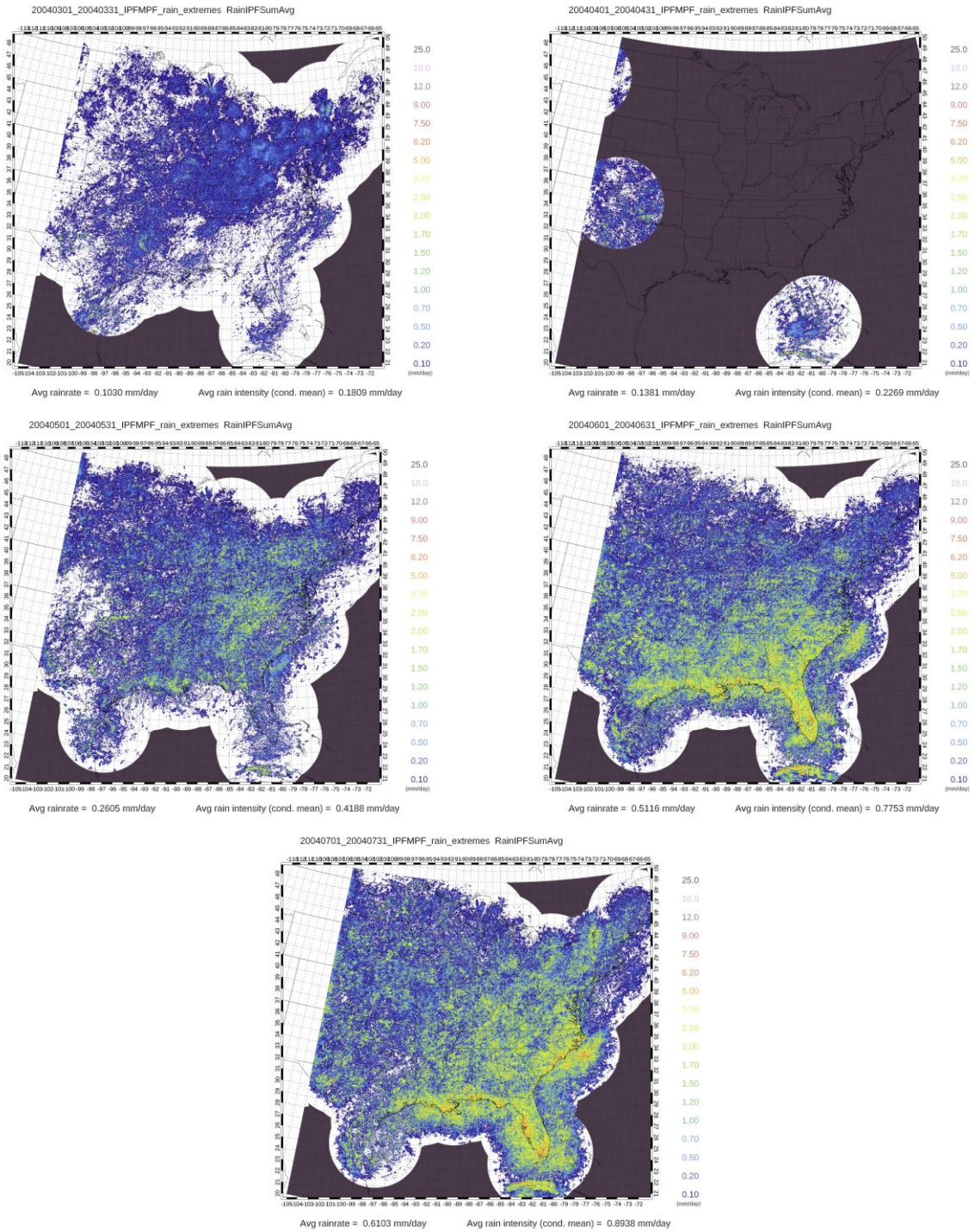


Figure 38: Monthly composite IPF rainfall maps for the year 2004.

panhandle. When looking at the monthly IPF precipitation composite maps in Figure 38, there is a more extensive IPF presence for the month of May in this area than in that of 2003, which is consistent with the onset map. Across the Southern Coastal Plain states, the date of onset remains similar to that of the seven-year average, with about 1-3 pentads variance. However, when looking at the April IPF rainfall composite map, a clear issue is present. As given by the black buffer covering most of the domain, a large amount of radar coverage is missing for the month of April, despite the fact that the corresponding data files exist. This is not a common phenomenon throughout the remaining months and years of the analysis. Since the observed hourly and daily and calculated pentad rainfall data only include the areas where valid radar coverage was present, it is likely that this may have played a role in there being no onset before pentad 25, or the beginning of May.

4.3.2.3 2005

For 2005, about 40% of boxes have observed onset, limited primarily to the SE US eastern-Central Plains, and the southern portions of Michigan and New York. All boxes are common to those associated with the average seven-year onset. The geographic extent of onset is comparable to that of the seven-year average, with onset extending from northward from South Florida into South New York and westward from the Carolinas into eastern reaches of the central plains (Figure 39). The earliest observed onset occurs in early-mid May in a broad region encompassing east Alabama, Georgia, South Carolina, and all of North Carolina except for the northeast quarter of the state. The next occurrences of onset occur southward in Florida starting in mid-May and progress into early June in the panhandle. Onset occurs in early June across the Ohio Valley, Gulf Coast and eastern central plains. Rather than observing the typical northward progression of IPF onset, onset seems to progress outward in all directions from this broad region

of first occurrence. Aside from this major difference, onset for 2005 is similar to that of the seven-year average. In most boxes where onset occurs in both 2005 and for the seven-year average, onset only varies in magnitude by 0-1 pentads. The lone exception to this is in South Florida (box 38), where onset in 2005 is 5 pentads later than that of the seven-year average in the same box. A point of interest is that there is no determinable IPF onset for box 35 along the Gulf Coast with this study's criteria for onset determination, despite fairly intense monthly IPF precipitation in the months of June and July (Figure 40). Nonetheless, the monthly composites of IPF rainfall for 2005 do well in supporting the observed dates of onset across the domain.

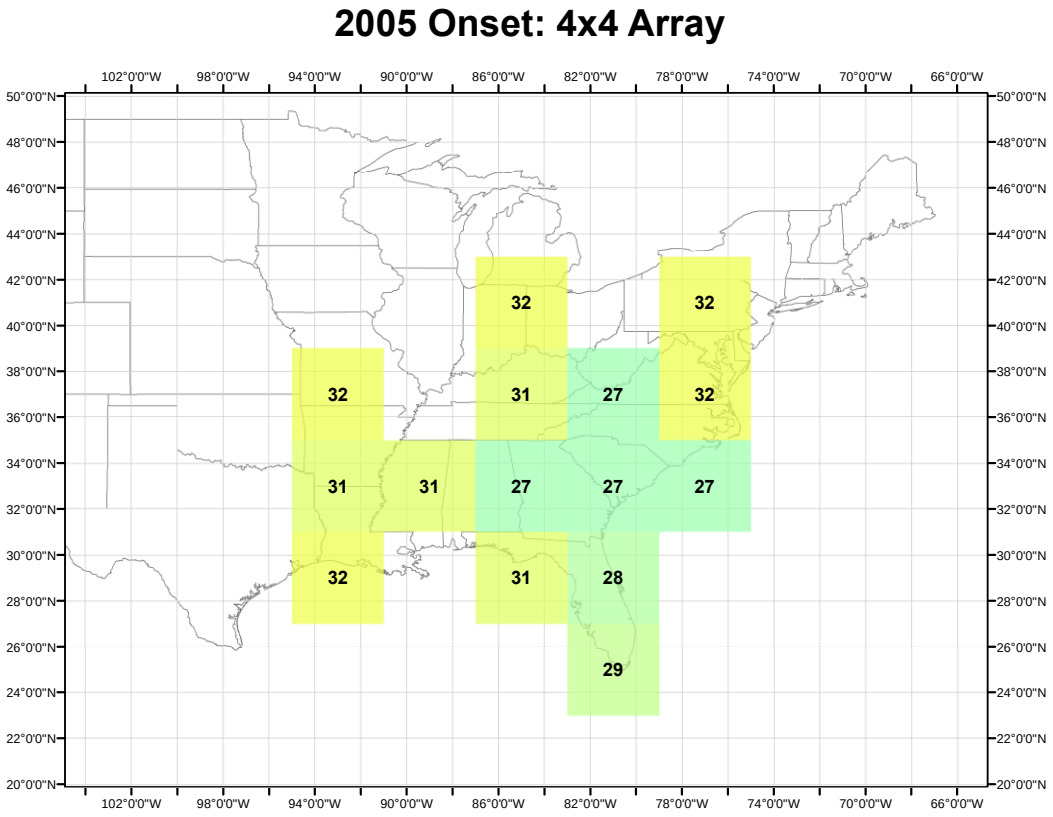


Figure 39: IPF Onset map for the year 2005. Domain-average onset occurs at pentad 32.

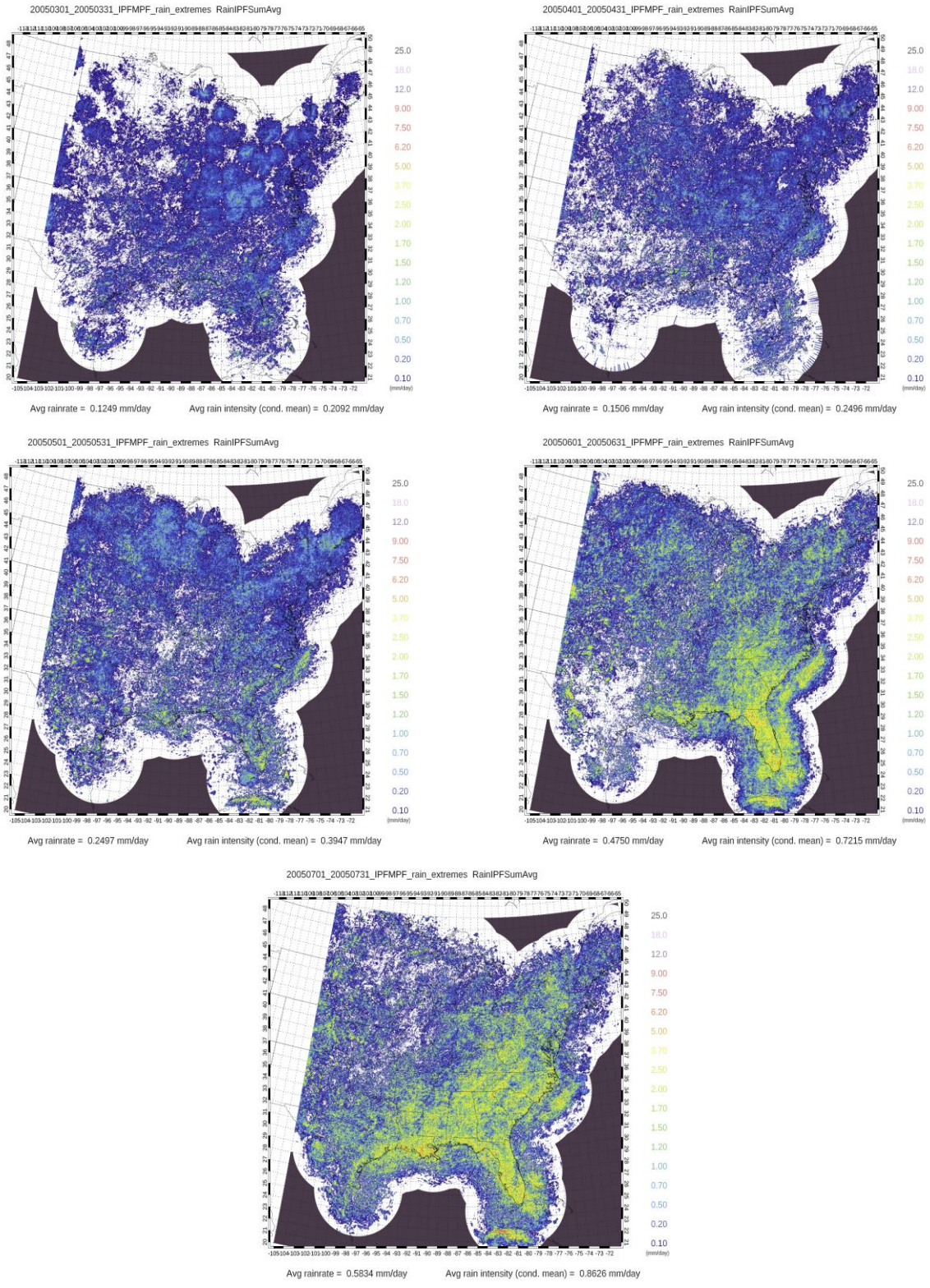


Figure 40: Monthly composite IPF rainfall maps for the year 2005.

4.3.2.4 2006

For 2006, about 40% of all boxes had determinable onset with the criteria used for this study, extending from the southeast quarter of Texas, into the SE US, and as far north as Michigan, northern Ohio and New York. Of these boxes with determinable onset, all are common with that of the onset domain for the seven-year average. The first occurrence of onset for the domain is at box 14 on pentad 27 (Figure 41). Along the Southeast Coastal Plain, onset progresses northward with time much like that of the seven-year average. South Florida (box 38) has an onset date of pentad 29 (May 21-25). IPF onset pushes northward along the coast over the course of the end of May (pentad 30) in North Florida and northeast North Carolina, the beginning of June in western North Carolina (pentad 31), and finally the back half of June

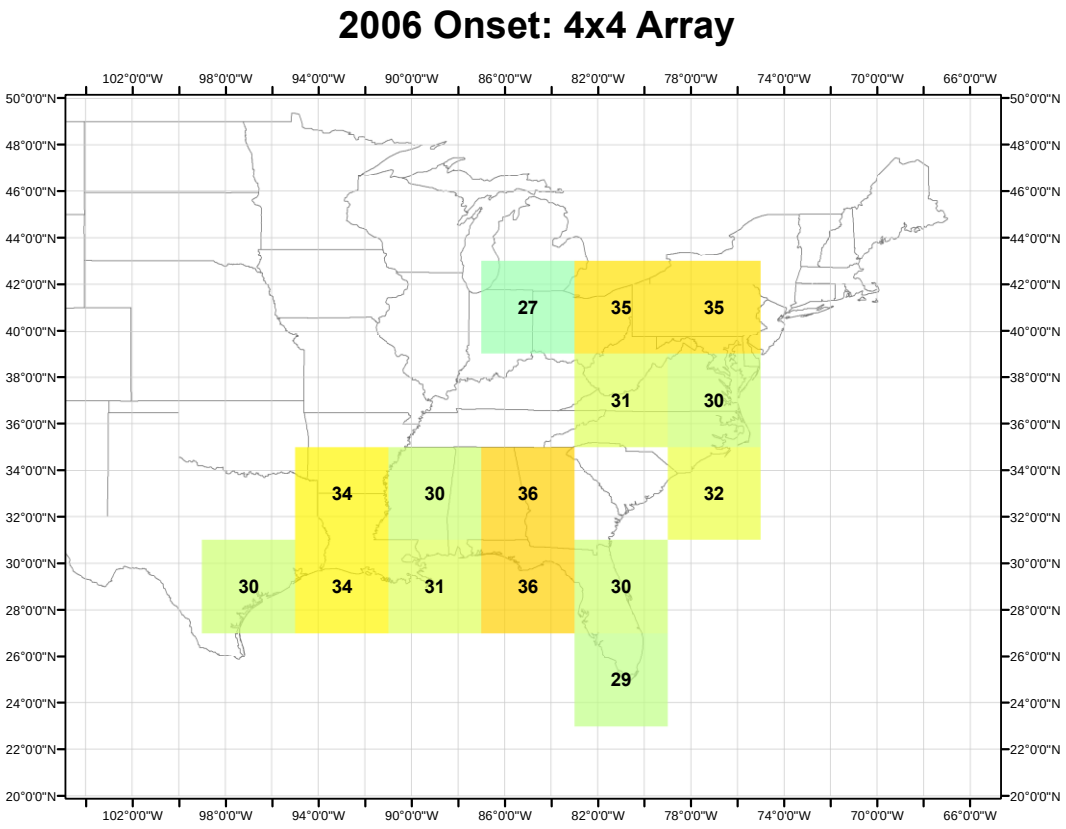


Figure 41: IPF Onset map for the year 2006. Domain-average onset occurs at pentad 31.

(pentad 35) in Pennsylvania, southern New York and eastern Ohio. This northward progression is also observed in the monthly IPF rainfall composites (Figure 42), particularly when looking at

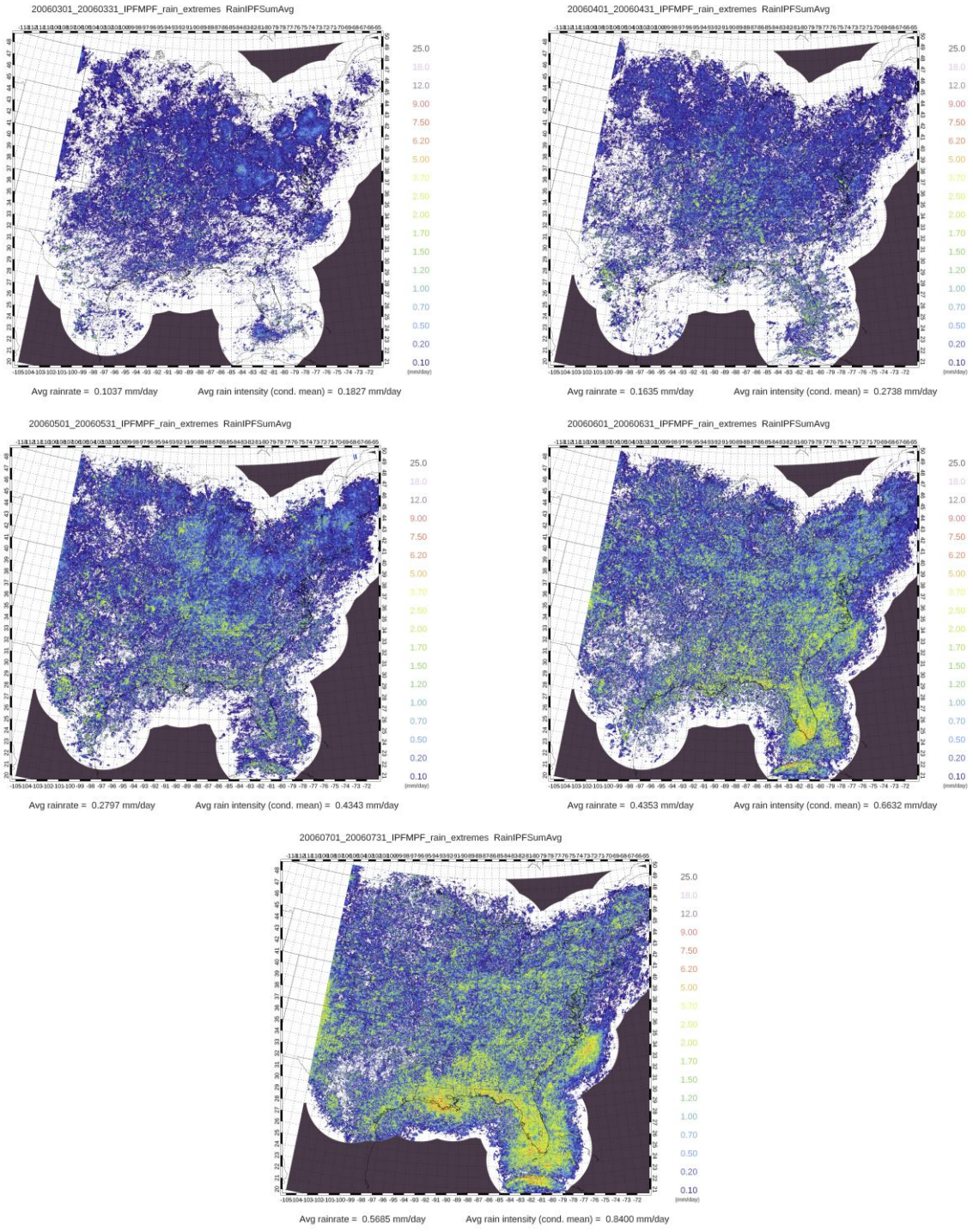


Figure 42: Monthly composite IPF rainfall maps for the year 2006.

the transition from May to June. When looking at the states of the Southern Coastal Plain, Gulf Coast and eastern portion of the central plains, the pattern of onset is not as simple to describe. In the Florida panhandle region (box 36) and just northward in west Georgia/east Alabama (box 29), the area has the latest onset occurrence of any box across the domain in this year, occurring at pentad 36. The boxes immediately west of this area (boxes 28 and 35) experience onset 5-6 pentads earlier. Conversely, as westward progression continues, boxes 27 and 34 have onset dates that occur 3-4 pentads later. As such, it is difficult to categorize a pattern of onset for this region of the domain, unlike with the Southeast Coastal Plain that more closely resembles the seven-year average with respect to pattern of onset progression.

4.3.2.5 2007

In 2007, onset occurs in almost half of the domain boxes, while only 12 are common to that of the seven-year average (Figure 43). Onset for 2007 is limited to the southern portion of

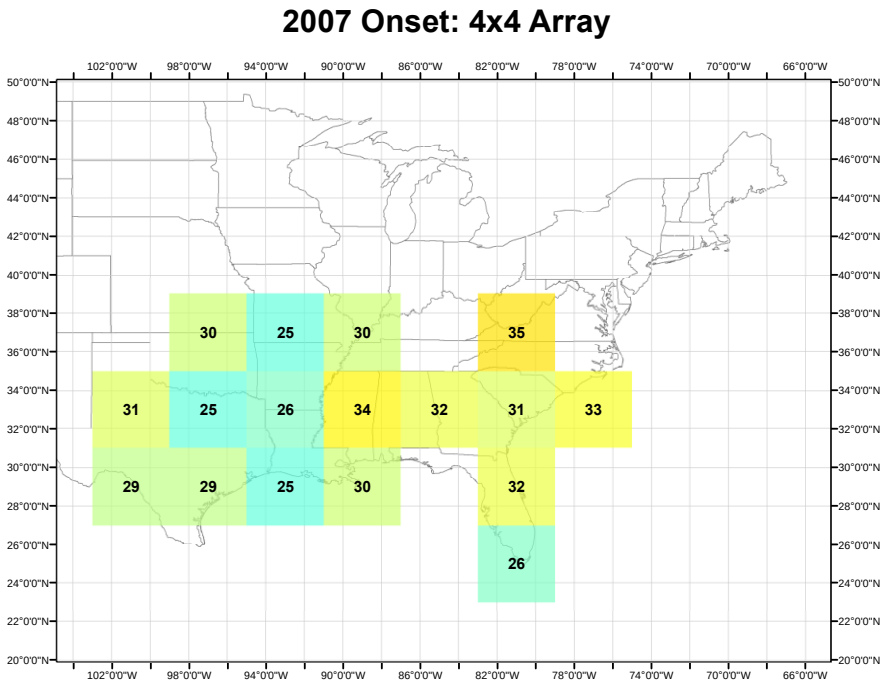


Figure 43: IPF Onset map for the year 2007. Domain-average onset occurs at pentad 31.

the Central Plains and the SE US south of central West Virginia. Up to this point, the onset domain for 2007 is the most unique of any prior year in the analysis. Along the Atlantic and Gulf Coasts onset progresses northward, beginning in South Florida on pentad 26 (May 6-10) and ending in southeastern North Carolina at pentad 33 (June 10-14). Unlike the years preceding 2007 in the analysis, no onset occurs north of box 31 along the Southeast Coastal Plain. Along the Gulf Coast, we also find the second instance where onset is not present in the Florida panhandle under the onset criteria of this study, following 2003 as mentioned in a preceding section. Further inland in the Carolinas, onset at box 23 is observed to occur at pentad 35 (June 20-24) and serves as the farthest north occurrence of onset for the South and Mid Atlantic. When looking at the lower Ohio Valley, western Southern Coastal Plain, and central plains' states, it is somewhat difficult to categorize the pattern of onset. Much like in 2006, as one progresses westward, onset fluctuates between a region of later occurrence, followed by a region of earlier occurrence, and finishing with another region where onset occurs later. In the case of 2007, boxes 21 and 28 are observed to have onset occurrence between mid-May and early-June. Onset in northeast Texas, eastern Louisiana, Arkansas and southern Missouri (boxes 20, 26, 27, 34) occurs at pentads 25 and 26 (Figure 44), with pentad 25 being the earliest observed date of onset for the domain in 2007. The dates of onset in this region of the central plains are interesting as it marks a 6-7 pentad average difference from the seven-year average. Further, the pentad 25 onset date for box 26 is unique in that this box does not have onset occurrence for the seven-year average. 2007 also marks the farthest west that onset has occurred up to this point in the analysis, with onset occurrence in western Texas at boxes 25 and 32. Interestingly enough, the monthly IPF rainfall composite maps for 2007 support these onset results (Figure 45). For the months of May, June and July, there is a far greater concentration of heavier average IPF rainfall across

Texas and Oklahoma than in any year preceding 2007. In particular, the average IPF rainfall in June for eastern Texas and Oklahoma rivals that of South Florida. Further, these findings are supported when looking at both the time series of boxes 26 and 27 (Figure 44) and looking at the spreadsheet of both observed daily and calculated pentad averaged rainfall. In the case of box 26, at pentad 25 there were three days above the 0.7638 mm onset threshold, with two days characterized by rainfall greater than 2.0 mm. At box 27, pentad 26 was characterized by three days with rainfall greater than the threshold, with two of the three being greater than 2.0 mm.

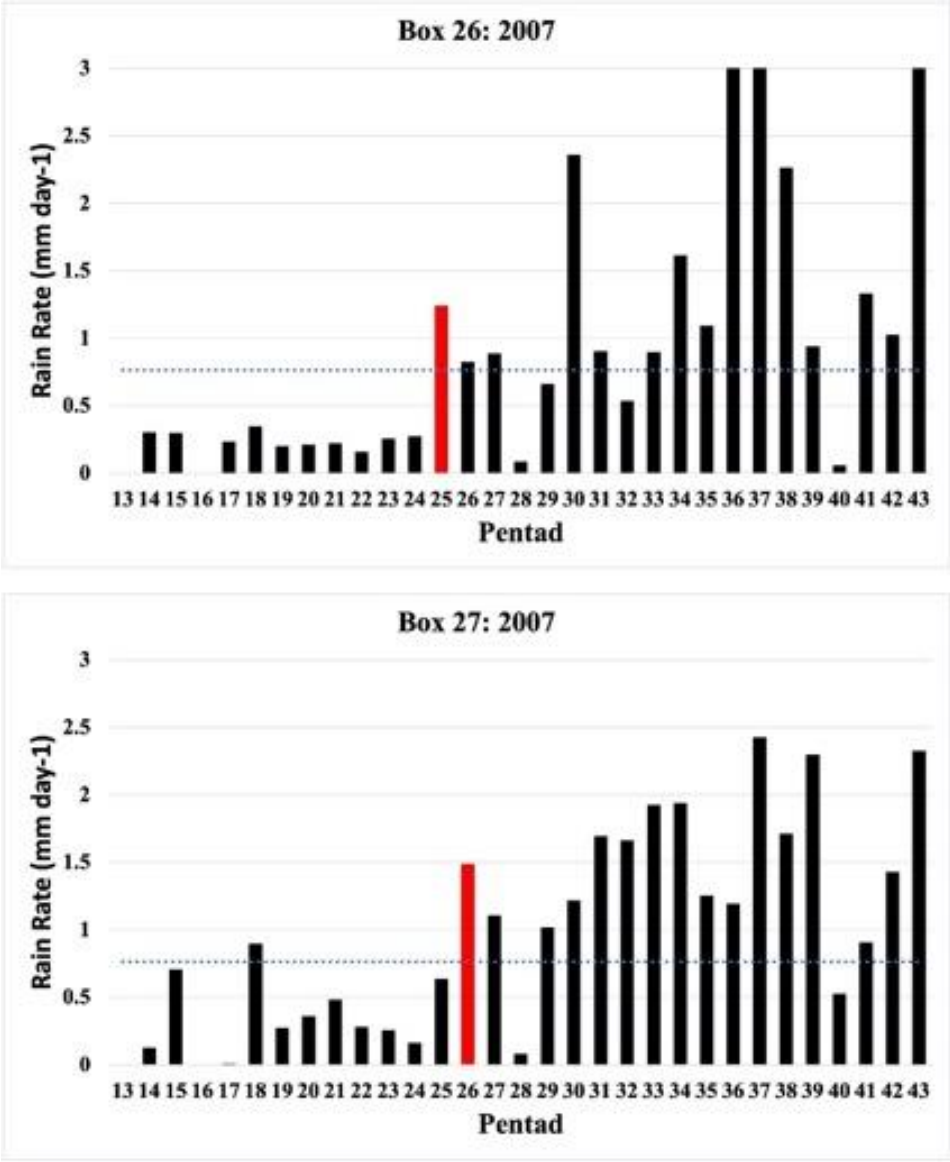


Figure 44: Time series for boxes 26 and 27 in 2007.

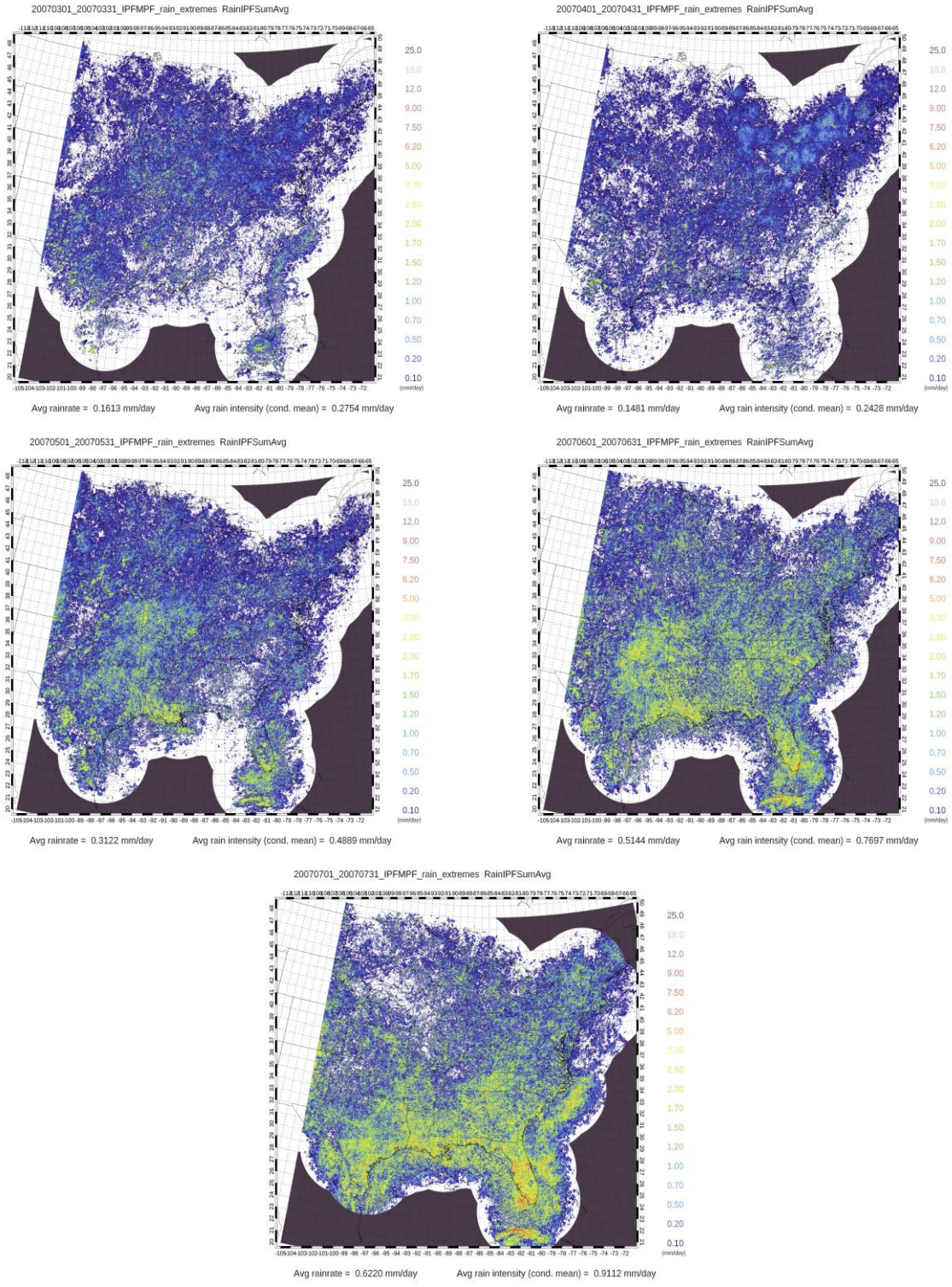


Figure 45: Monthly composite IPF rainfall maps for the year 2007.

4.3.2.6 2008

Onset occurs in a little more than a quarter of boxes for the year 2008, all of which are common to that of the seven-year average and limited to the SE US (Figure 46). The earliest occurrence of onset occurs at pentad 18 (box 27). 2008 presents another unique year for the onset analysis, as missing data during the months of April and June have played a role in the patterns of both onset timing and spatial extent of onset. This is apparently clear when looking at both the onset map (Figure 46) and IPF monthly rainfall composite maps (Figure 47). With the exception of box 27, the dates of onset for the entire onset domain have little to no variation from box to box. Likewise, for the IPF rainfall composite maps, both an extreme uptick in IPF rainfall from June to July in key areas, in particular South Florida and the Gulf Coast, and extreme spottiness

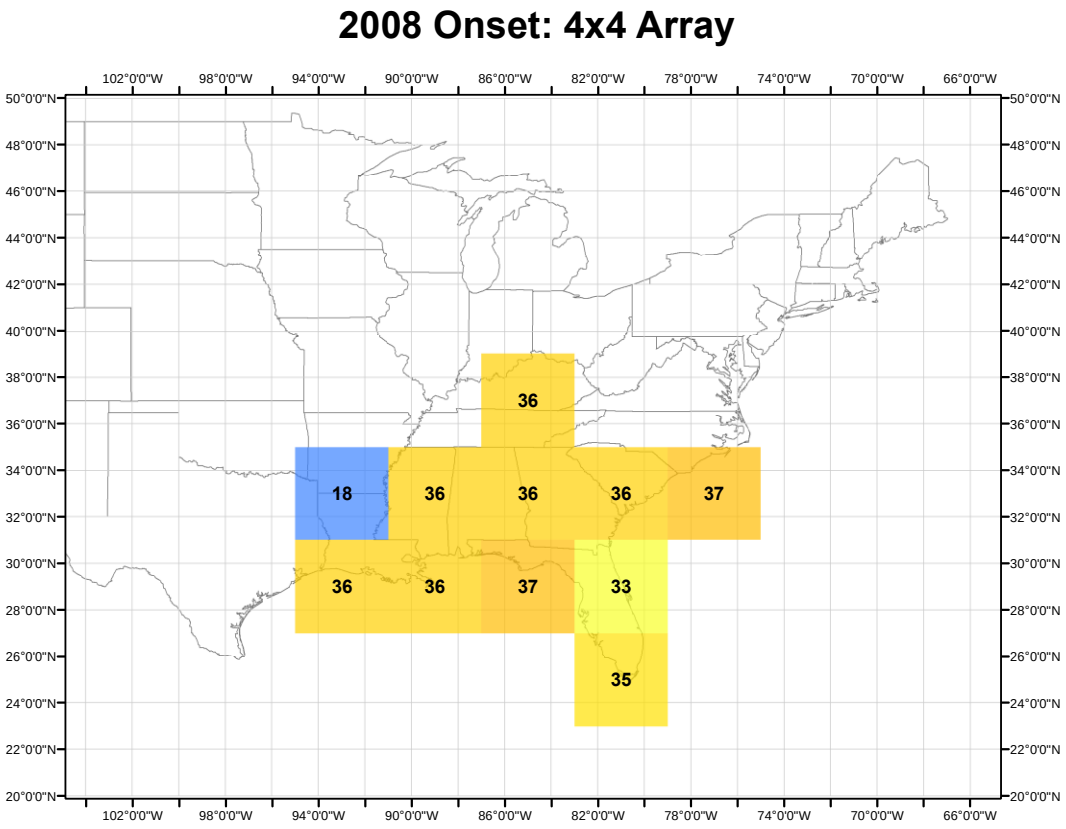


Figure 46: IPF Onset map for the year 2008. Domain-average onset occurs at pentad 36.

of the April and June maps suggests that many of the original instantaneous hourly rainfall data files are missing. As such, both the spatial and temporal patterns of onset observed for 2008 in this study are not accurate.

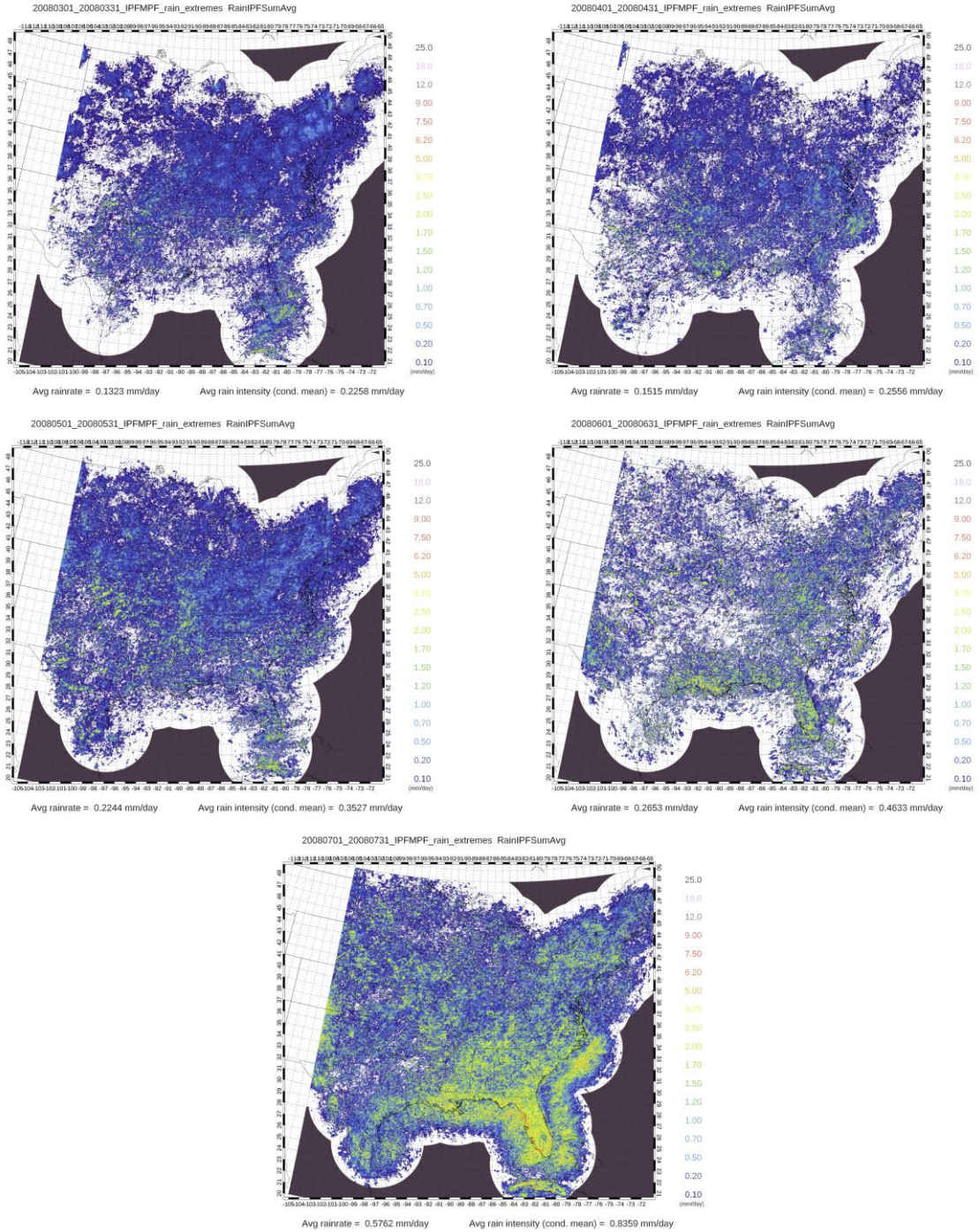


Figure 47: Monthly composite IPF rainfall maps for the year 2008.

4.3.2.7 2009

For 2009, a little more than a third of boxes have IPF onset occurrence, with 13 being common with the seven-year average onset domain (Figure 48). Onset for 2009 is common across the SE US and into eastern-Ohio/western-Pennsylvania, but onset also occurs in two boxes south of the Texas panhandle and in central Nebraska. Across most of the Southeast Coastal Plain region of the SE US, onset generally occurs in early-mid May at pentad 25-26, though in box 30 (South Carolina coastal plain) onset occurs much earlier in mid-April (pentad 21). Onset for the eastern half of Florida and coastal North Carolina only varies by 2 pentads, ranging from the first pentad of May (pentad 25: May 1-5) and into mid-May (pentad 27: May 11-15). About 8% of hourly instantaneous rainfall data files are missing for April of 2009 between the 1st and 23rd days of the month. The 2009 monthly IPF rainfall composite map

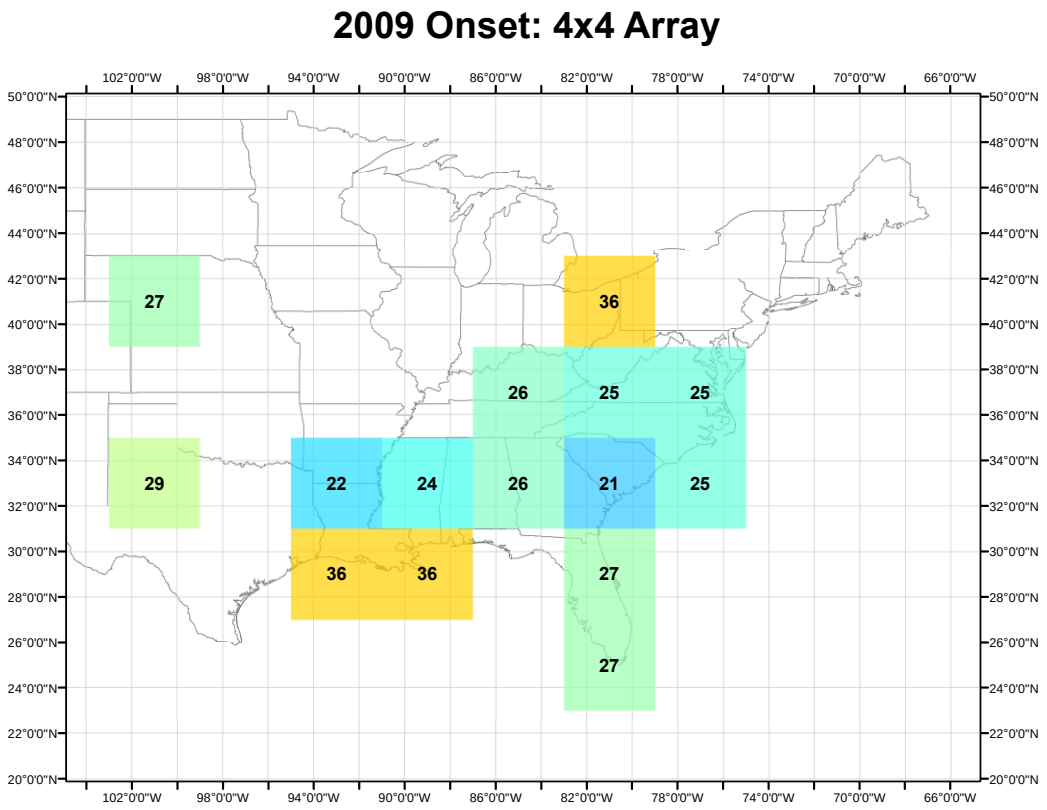


Figure 48: IPF Onset map for the year 2009. Domain-average onset occurs at pentad 29.

(Figure 49) doesn't suggest much difference from average for the month of April, so this suggests that the aforementioned missing or corrupt hourly rain data has influenced the observed results. There are two other boxes that have observed dates of onset in the month of April: boxes

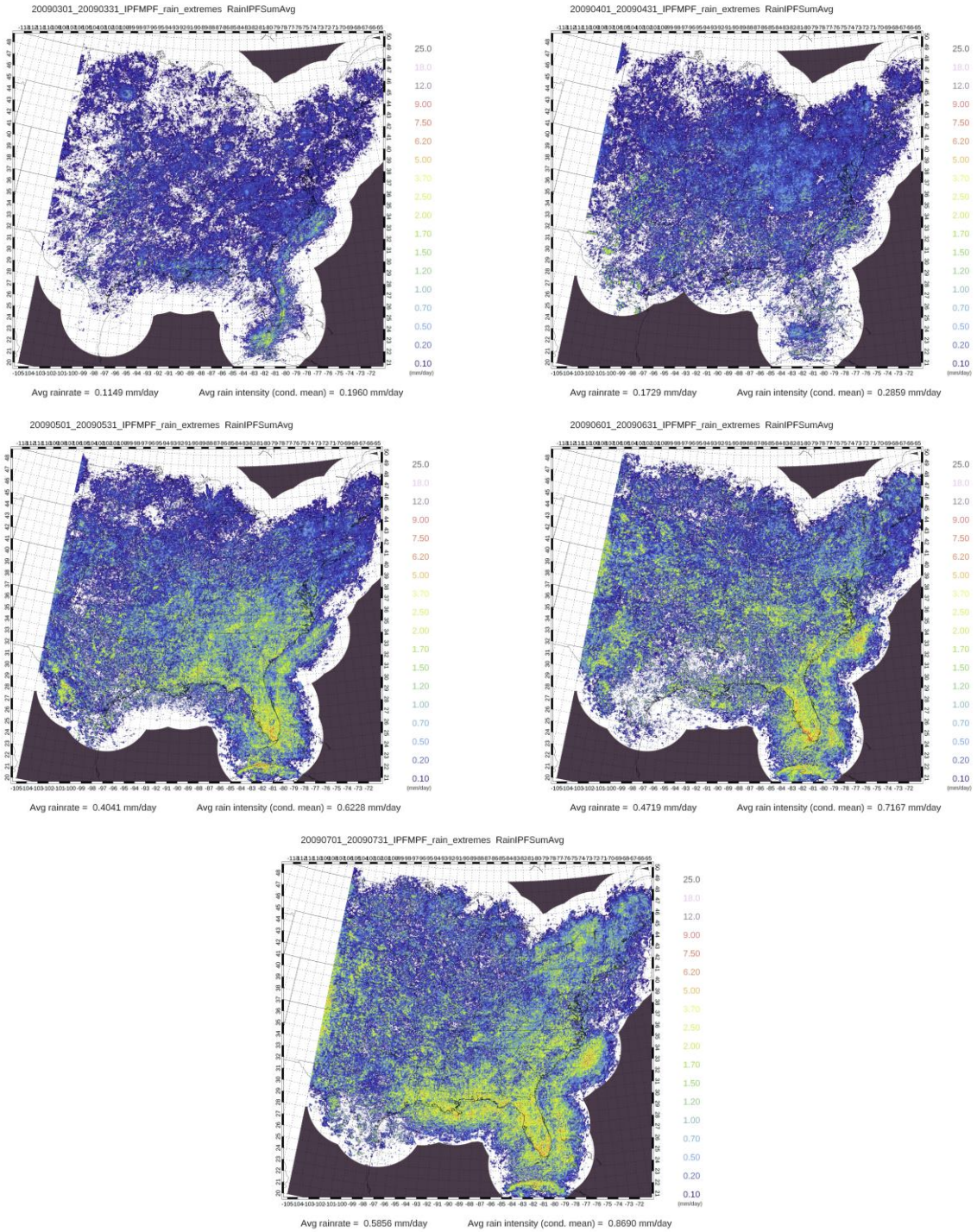


Figure 49: Monthly composite IPF rainfall maps for the year 2009.

27 and 28 (northern Louisiana, southern Arkansas, Mississippi, western Alabama). These two onset dates are early for these two individual boxes, as onset for these boxes occurs 10 and 7 pentads earlier than average in 2009. While noticeable that there are some rainfall recording issues for the months of May and June at these boxes, this does not appear to be the case for April (Figure 50). In the cases of boxes 27 and 28, the earlier than normal observed onset is due to intense rainfall over the course of a few days or a single day. For Box 27, 2.7123 mm of rain was recorded for April 18th, falling in pentad 22. For box 28, 1.942 mm, 2.127 mm, and 2.866 mm of rain were measured during the three-day period of April 28-20, or the final three days of

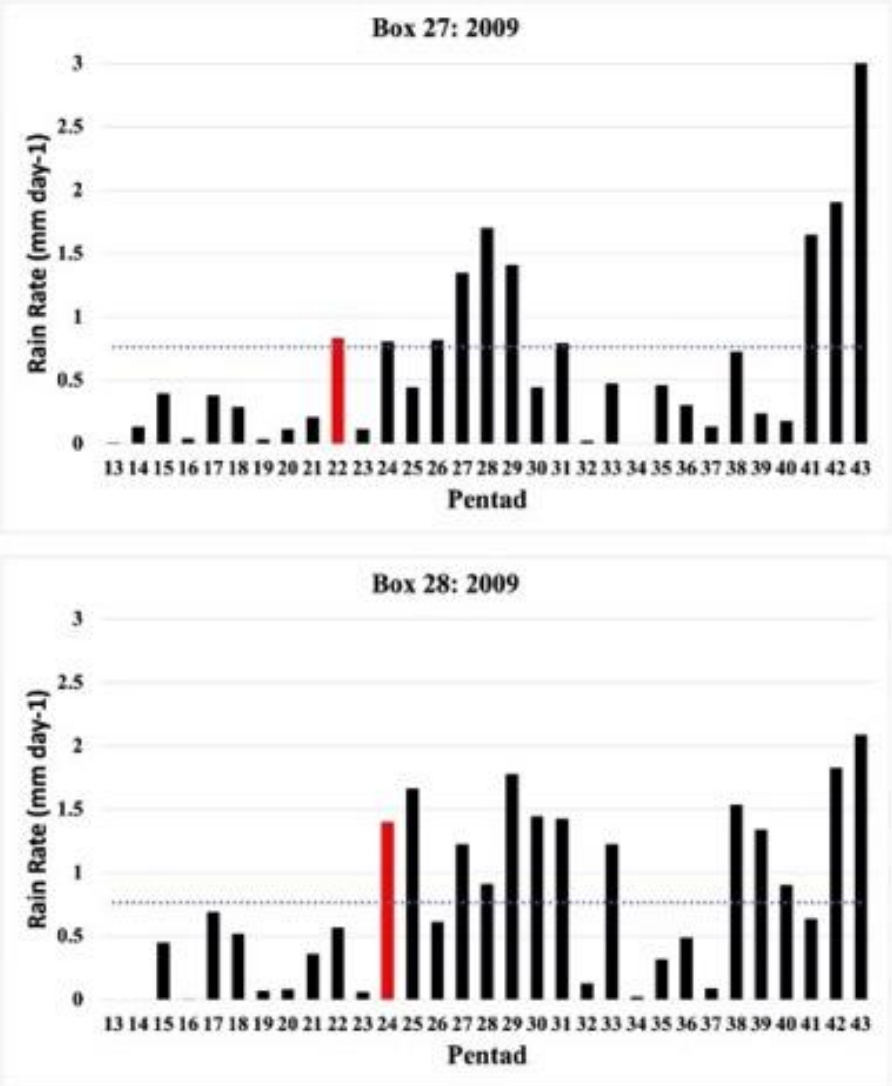


Figure 50: Time series for boxes 27 and 28 in 2009.

pentad 24. Onset in the two boxes directly south (boxes 34 and 35; southern Louisiana and Mississippi) occurs at pentad 36, which is 14 and 10 pentads later than that of boxes 27 and 28, respectively. As one continues to look westward, two isolated occurrences of IPF onset are present in 2009. These isolated cases occur at boxes 10 and 25 and occur in mid-to-late May (pentads 27 and 29). As such, it is difficult to speak of any particular pattern in onset timing for 2009. While slightly earlier than average, the timing and pattern of onset along the Southeast Coastal Plain is not all that abnormal. However, onset west of boxes 22 and 29 becomes difficult to identify patterns for. Regarding spatial extent, the extent of the broader area of onset occurring boxes of the SE US is about what one would expect compared to average. The two isolated cases in the central plains, however, is a phenomenon not typically experienced in this analysis. However, as mentioned previously, missing or corrupted hourly rain data influences the analysis results, and as such corrupts the 2009 onset analysis.

4.3.2.8 2010

In 2010, one-half of all boxes across the domain have determinable onset, with all but 3 being common to the onset domain of the seven-year average (Figure 51). With the exception of the southeast quarter of Texas and eastern Nebraska, onset is limited to the SE US and a small, northern portion of the domain in Illinois, Indiana and western Ohio. Across the SE US west of boxes 21, 28 and 35, onset timing is similar. Beginning in south Florida and progressing northward through the lower Ohio Valley and the mid-Atlantic, onset only varies by 2 pentads in the early and middle portions of May (pentads 26-28). Though more abruptly than other years, onset still progresses northward with time across the SE US in this case. Onset at and west of boxes 13, 21, 28 and 35 is a bit different than that of the SE US. This area encompassing the western Gulf Coast and central plains has onset date variance in the magnitude of 13 pentads

from earlies to latest occurrence. Earliest occurrence happens at pentad 23 (April 21-25) at boxes 21 and 28. The latest case of onset occurs in the southeast quarter of Texas (box 33) at pentad 36 (June 25-29), meaning that there is greater than a 2-month difference in the difference of onset timing from earliest to latest occurrence in this region. Interestingly enough, there exists a noticeable onset pattern for this region. Onset seems to progress outward from boxes 21 and 28, much like how onset often progresses northward from south Florida across the SE US. In this case though, onset progresses in all directions from these two boxes, including eastward into the previously described regions of the SE US west of east of boxes 21, 28 and 35. Onset progression in this manner is more gradual than that of the traditional northward progression of the SE US, with onset varying by 4-9 pentads at the boxes bordering boxes 21 and 28. Further, there is an untraditionally strong IPF rainfall signal for the month of May (Figure 52) in Louisiana and northward into Arkansas, Mississippi, and western Tennessee. The regions of Mississippi and Tennessee are of particular interest because of the abnormally early date of onset

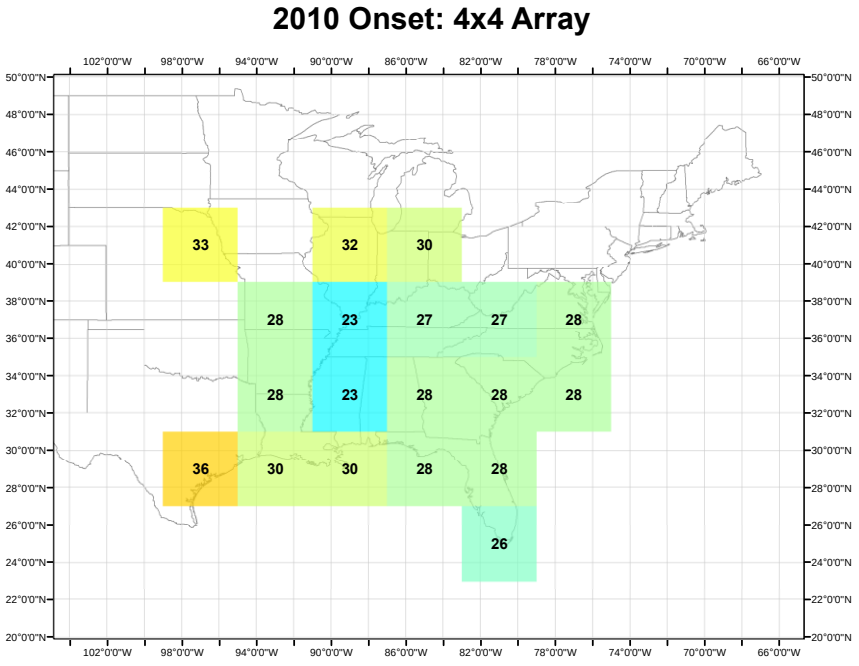


Figure 51: IPF Onset map for the year 2010. Domain-average onset occurs at pentad 30.

in these areas (boxes 21 and 28, specifically). Moving into June and July, the traditional trend of expansive extent and magnitude of IPF rainfall is observed across the entire SE US, with the most intense IPF signals across the Gulf Coast, south Florida and offshore of North Carolina.

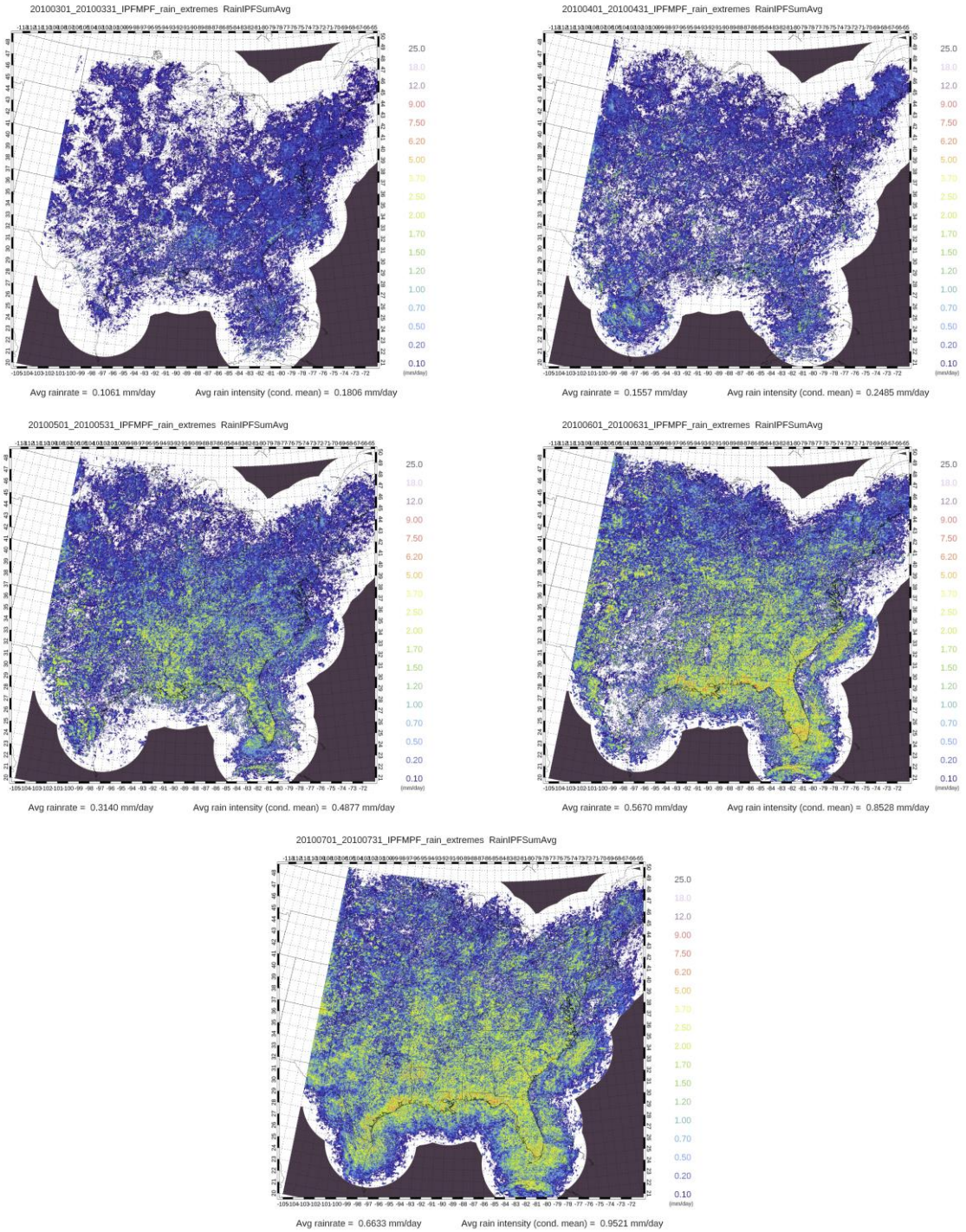


Figure 52: Monthly composite IPF rainfall maps for the year 2010.

4.3.2.9 2011

In 2011, determinable onset occurs at one-third of all boxes, all of which are common to the onset domain of the seven-year average (Figure 53). Though each of the 13 boxes with observed onset are common to that of the seven-year average, northward progression of onset is not observed in 2011. The earliest occurrence of onset is pentad 27 (May 11-15) at boxes 23 and 24, encompassing most of North Carolina, Virginia, and the southern portions of West Virginia and Maryland. The onset domain only extends north of this area by one box (box 16). Otherwise, onset progresses westward into the Ohio Valley and southward and southwestward into the states of the Atlantic and Gulf Coasts from these two boxes, with the difference in earliest and latest onset occurrence being 8 pentads (pentad 35 is the latest observed occurrence of onset and is in the Florida panhandle). Onset across this broad region occurs between pentads 30 and 35, or the

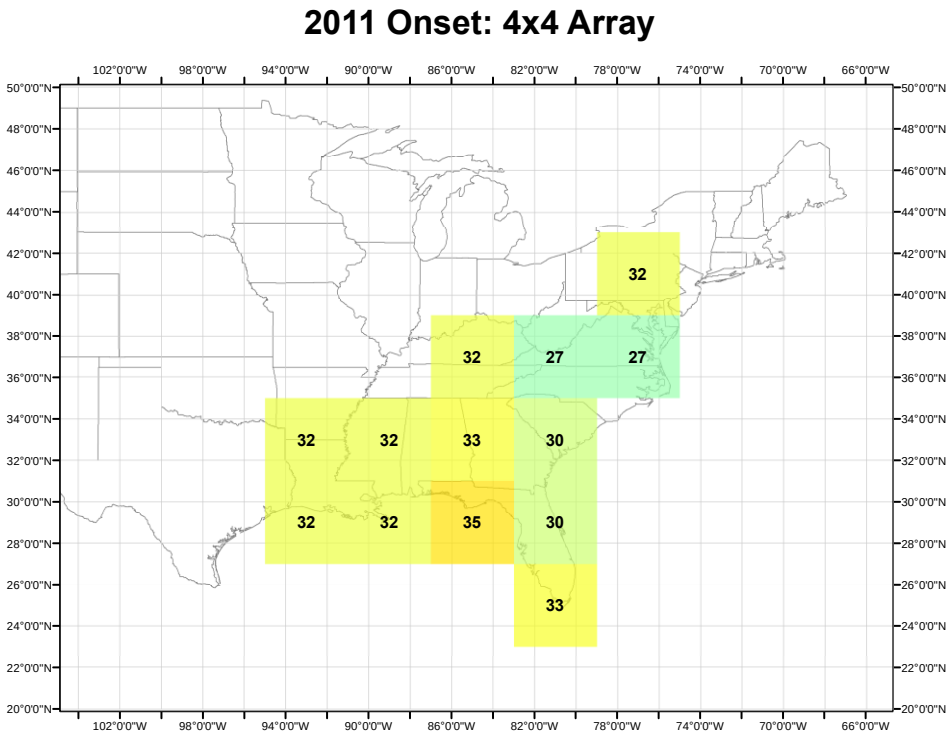


Figure 53: IPF Onset map for the year 2011. Domain-average onset occurs at pentad 33.

approximately month-long period between the dates of May 26 and June 24. In essence, this mostly southward progression of the IPF onset season is the opposite of what is observed for the seven-year average onset analysis. When looking at the monthly IPF rainfall composite maps for 2011 (Figure 54), the IPF signatures for the months of April and May are very spotty across

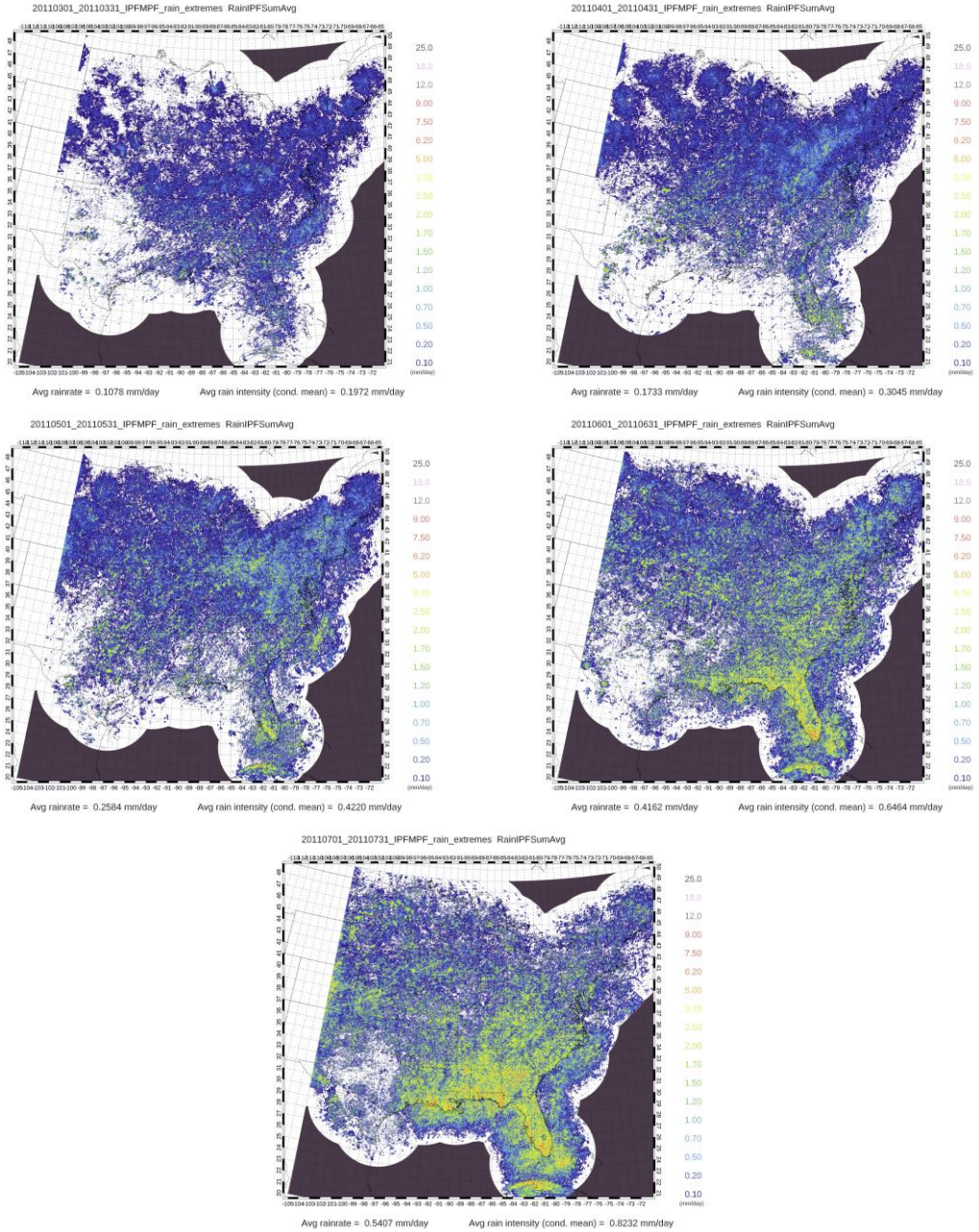


Figure 54: Monthly composite IPF rainfall maps for the year 2011.

many of the Southern Coastal Plain and Gulf Coast states, including the Florida panhandle. This is not the case in the area within boxes 23 and 24, as the IPF signature for this analysis region is steady when moving from April to May. June and July are characterized by the characteristic uptick in IPF precipitation across most of the SE US, but the earlier sustained IPF signature of boxes 23 and 24 is a likely reason for the earliest observed onset occurring in this area.

4.3.2.10 Summary of Year-To-Year Differences

While many of the overarching trends in onset are similar from year-to-year, there are some key differences to discuss as well. The most obvious, but also most powerful, is that onset at each individual box changes from year-to-year (Figure 55). With standard deviation, like that shown in Figure 55, the closer a value is to zero indicates that the date of onset at each individual box is similar from year-to-year. Likewise, the further a value is from zero is indicative of

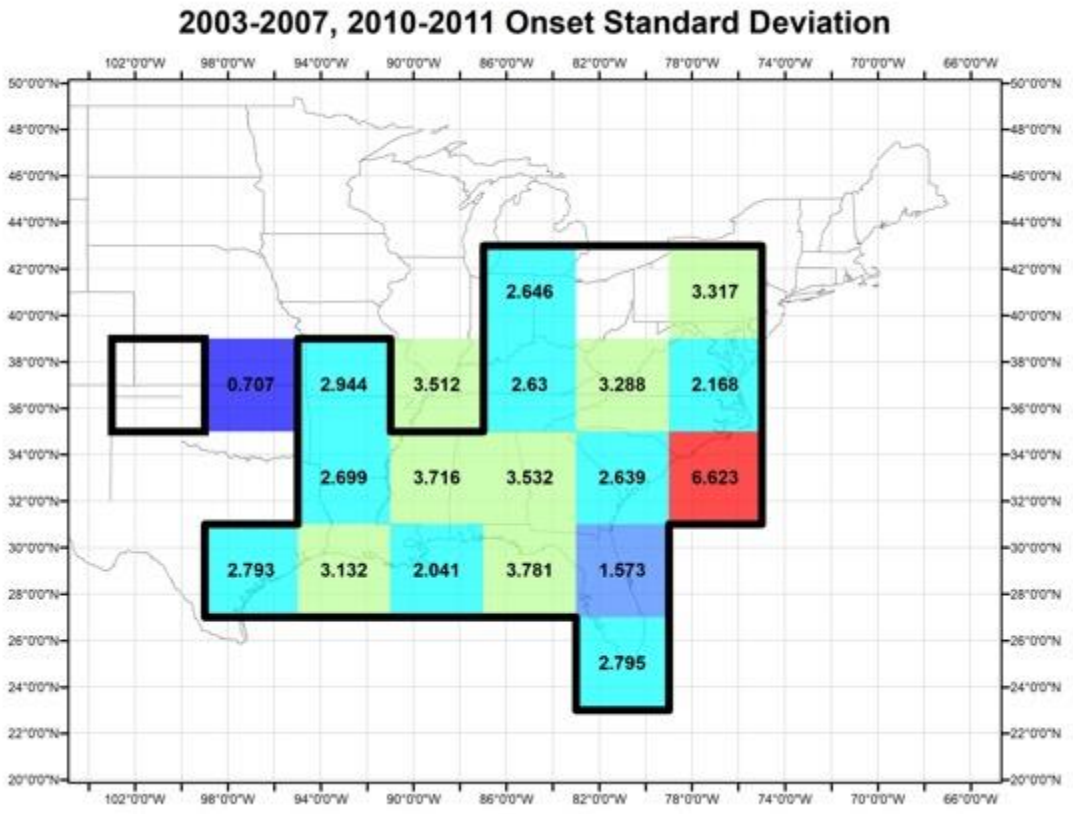


Figure 55: Same as Figure 34, but for the standard deviation.

greater year-to-year difference in onset date. For this analysis, of the 19 boxes included in the seven-year average onset domain, 17 had multiple years where onset occurred. As such, it was possible to calculate their standard deviation. Of the boxes included within the seven-year average onset domain, outlined in black, the least year-to-year difference in onset timing was observed at box 37 in northern Florida. Onset occurred at box 37 in each of the seven years, with onset occurring at pentad 28 in four (2003, 2004, 2005, 2010) of the seven years. The greatest difference in year-to-year onset date was observed at box 31 in the southeastern North Carolina coast. Here, onset occurred in six out of seven years, with an 18-pentad difference between earliest occurrence (pentad 15 in 2003) and latest occurrence (pentad 33 in 2007). In two instances, boxes 15 and 18, the standard deviation could not be calculated because onset was only observed in a single year out of seven in these areas. However, because both of these boxes still had a large number of pentads where the rain rate met or exceeded the threshold value, the onset criteria for these two boxes were still met when the pentad IPF rainfall was averaged over the seven-year period. It should be noted, however, that box 15 would have had a measurable standard deviation in the event that 2008 and 2009 were included in the calculation of the standard deviation. This is because box 15 had observed onset at pentad 36 in 2009.

As previously noted, there were a number of years (2003, 2007, 2009) where onset did not occur in the western Florida panhandle (Figure 56). For each of these three years, IPF rainfall evolution differs from that of the seven-year average at box 36 (Figure 57). At box 36, IPF rainfall tends to progressively increase beginning at pentad 28, though the first most noticeable jump occurs at pentad 31. In the case of 2003, the only dates where the rainfall appears invalid is at pentad 30, as May 26 and 28-30 are shown as null. However, this pentad alone would not have been enough to meet the onset criteria, and as such box 36 in 2003 is different from average. In

2007, portions of data had null precipitation values where the corresponding rainfall value did not exist for the following days: May 18-22 (pentads 28 & 29), May 24-May 29 (pentads 29 and 30) and June 15-17 (pentad 34). Much like 2003 however, it is unlikely that filling the gaps in

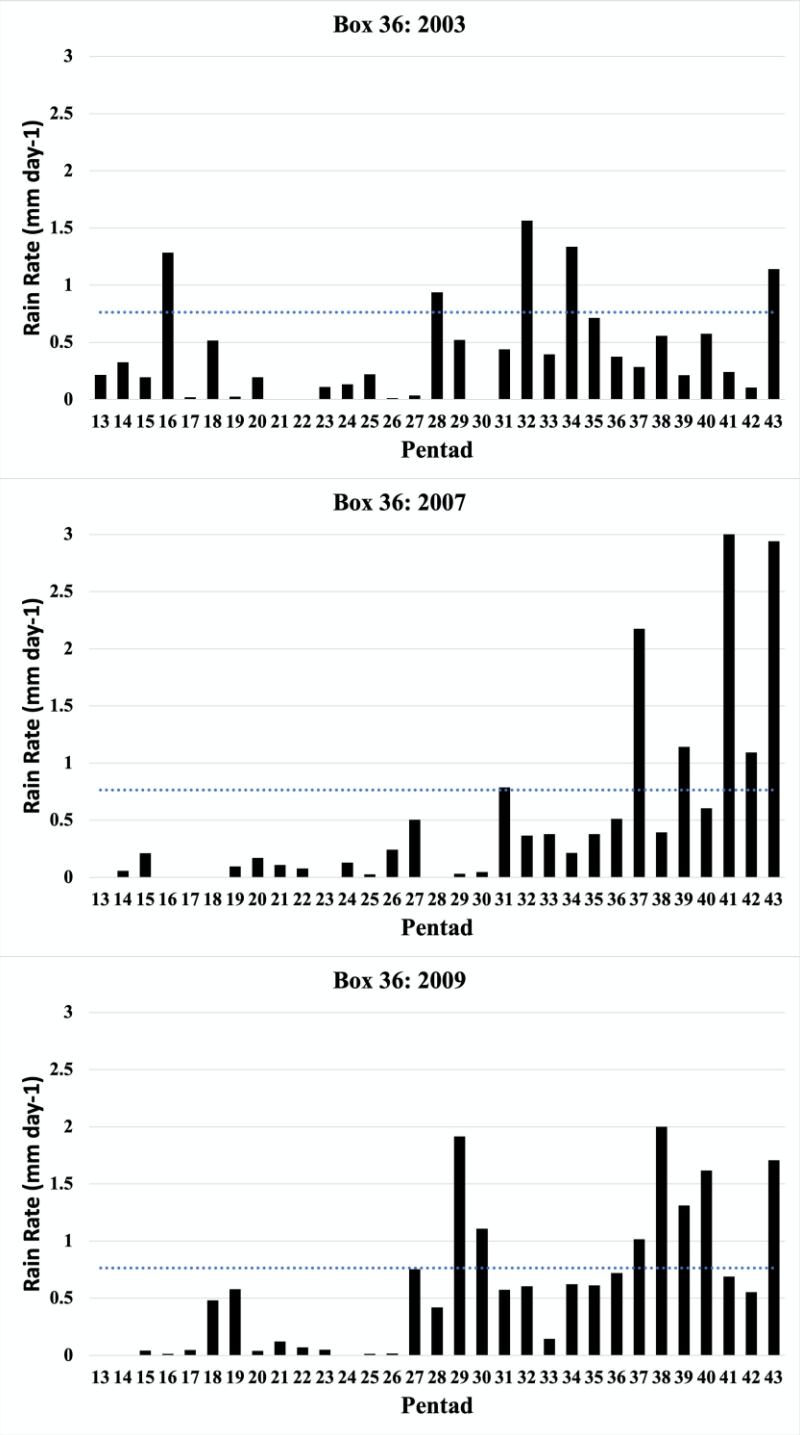


Figure 56: Time series of box36 in 2003, 2007 and 2009.

data from pentads 28-30 would be sufficient enough to support meeting the onset criteria at box 36 in 2007. This conclusion was made based on the fact that pentads 31-33 and beyond pentad 35 are complete. As such, the difference of IPF progression at box 36 in 2007 is not related to data quality. In 2009, over 31% of data is missing for the month of June between pentads 31 and 35 after 1400 UTC. Due to the majority of this missing precipitation data being in the early-midafternoon and early evening (the time of diurnal maximum of IPF rain in the summer), it is likely that the inclusion of this data would have increased the daily and pentad IPF rain rates to an intensity sufficient enough to meet the criteria of onset. As such, the abnormality of IPF rain progression for box 36 in 2009 can be considered as an issue of data quality.

4.4 Year-to-Year Spatial and Temporal Behavior of Onset

This section begins by describing the year-to-year spatial extent of onset. Following will be a discussion on the temporal behavior of onset from year-to-year. In an effort to both quantitatively and qualitatively compare year-to-year onset characteristics, Figure 58 shows each

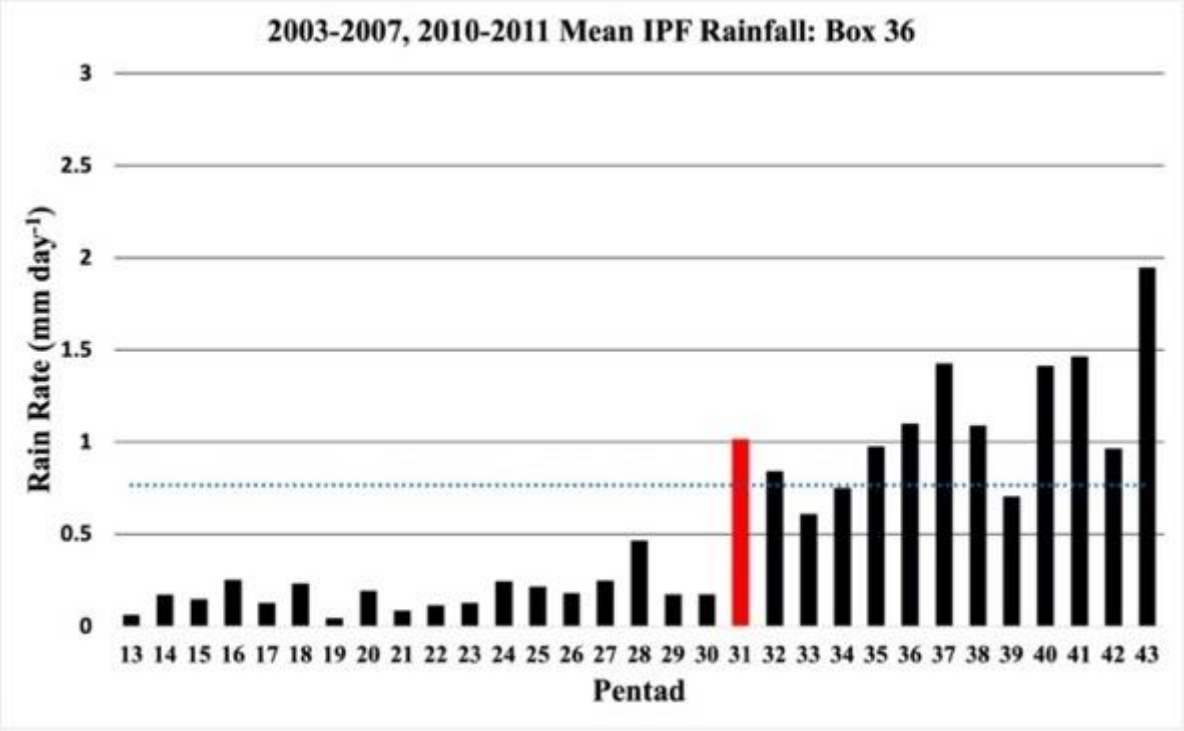


Figure 57: Seven-year average time series of IPF rainfall for box 36.

of the previously introduced onset maps for each of the nine years, in addition to that of the seven-year average.

4.4.1 Spatial Extent of Onset

From year-to-year, one main commonality exists when observing the spatial extent of IPF onset. This noticeable feature from year-to-year is the presence of onset in the SE US. In particular, onset occurs each year along the Southeast Coastal Plain from the Gulf Coast to the western border of the Florida panhandle, in addition to the inland reaches of the Southern Coastal Plain states of Louisiana, Mississippi, Alabama and Georgia. This broad region of year-to-year onset was also observed by Rickenbach et al. (2020) using an array of 2° x 2° boxes.

Not included in the list of common regions of this discussion are the following: North Carolina (boxes 23, 24, 30, 31), Tennessee (boxes 21 and 22), and the Florida panhandle (box 36). These regions mostly have year-to-year occurrence of onset when looking at the 2° x 2° box array of Rickenbach et al. (2020) for the years of 2009 through 2011. In this analysis, this is not always the case. For North Carolina, onset for one or more of boxes 23, 24, 30, or 31 was not present in 4 of the 9 years in the analysis (2006, 2007, 2008, 2011). Likewise, there are three occurrences where onset is not observed at box 36 for the Florida panhandle (2003, 2007, 2009), though 2007 and 2009 alike have data quality issues during the month of April in this region. Particularly in the boxes encompassing coastal North Carolina and the Florida panhandle, a possible explanation for this is the size of the analysis region itself. It is possible that onset for these boxes was not picked up with the onset criteria used in this analysis because a portion of the box area is situated over water. With the Rickenbach et al. (2020) 2° x 2° domain, the boxes associated with these regions had a greater portion of areal coverage inland. This will be

discussed further in Section 4.3 in a discussion related to onset sensitivity when comparing the onset results of 4° x 4° and 2° x 2° arrays used in this analysis.

The westernmost extent of the onset domain is extremely variable from year-to-year. In six of the nine analysis years (2003, 2005, 2006, 2008, 2010, 2011), onset is not present west of the column of boxes that encompasses boxes 2, 11, 19, 26 and 33 (southward of western Minnesota and the eastern Dakotas), meaning that onset remains present primarily in the SE US. Of these six years, the westernmost extent of the onset domain does not include these particular boxes on three occasions (2005, 2008, 2011). This leaves three years (2004, 2007, 2009) where the westernmost extent of the onset domain is beyond the aforementioned column of boxes. For the years of 2004 and 2007, these boxes are not isolated, but rather are part of a broad region of onset occurrence in the Central Plains. In 2009 the main region of IPF onset occurrence is across the SE US and east of Texas. However, two isolated cases of onset are present in the Central Plains at boxes 10 and 25. As previously mentioned, summer rain in the central US relies primarily on moisture transported from the Gulf of Mexico via the low-level jet (Carbone and Tuttle, 2008).

Onset does not occur in any of the years between 2003-2011 in the northern border of the domain (boxes 1-9). This does not come as a surprise for many of the same reasons as that of the central US. Just south of this area at boxes 13, 14, 15 and 16 (Northern Illinois to eastern Pennsylvania), onset occurs in 6 of the 9 (2004, 2005, 2006, 2009, 2010, 2011) analysis years. This phenomenon is also observed for the seven-year average onset domain at boxes 15 and 16. While not totally distinguishable on the onset maps, these boxes correspond to the vicinity of the

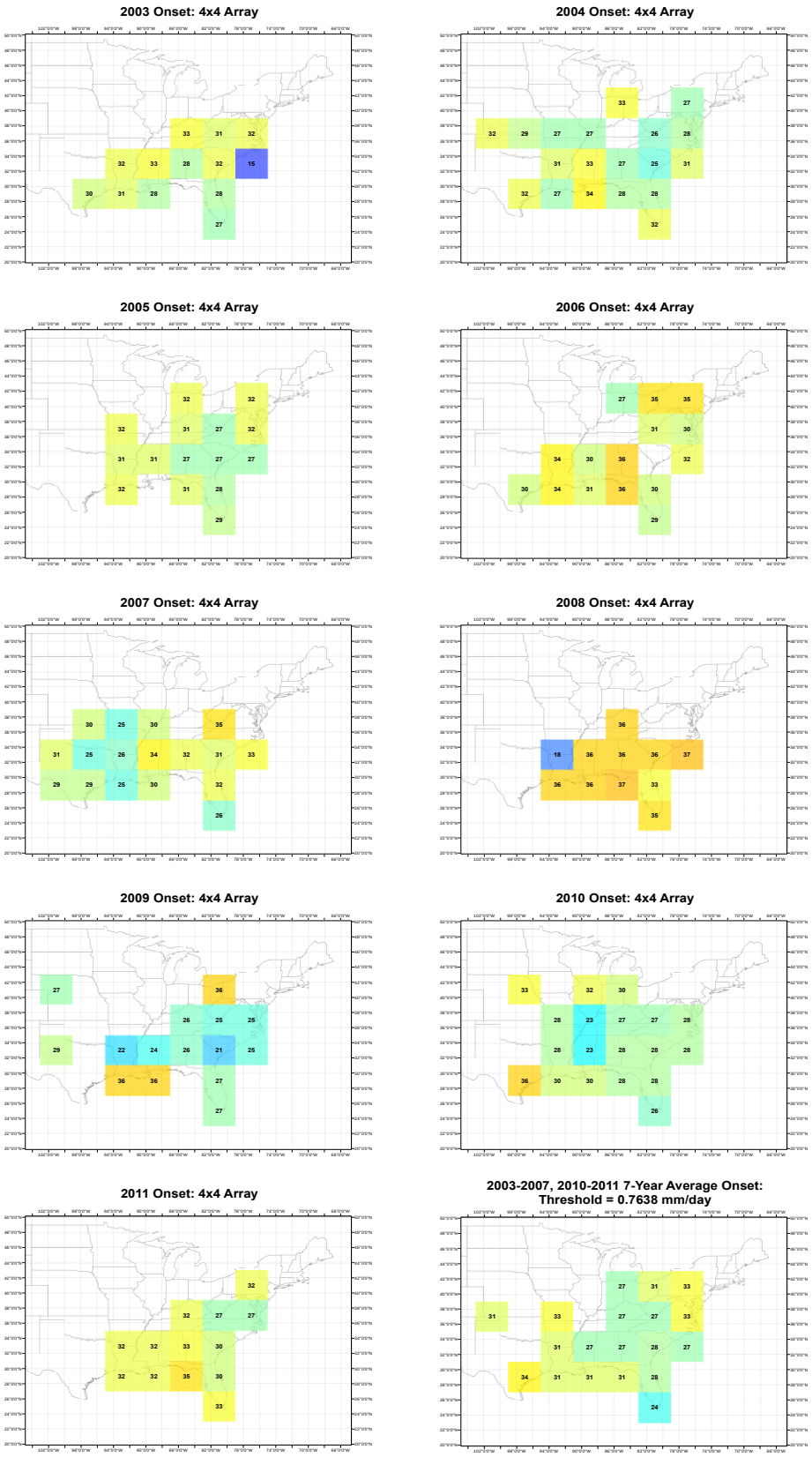


Figure 58: Year-to-year and seven-year average onset maps previously introduced in Section 4.1.

Great Lakes. While beyond the scope of this thesis, it potentially warrants further investigation whether the lake breezes off of these water bodies influence IPF onset similarly to that of the sea breeze circulation across the SE US, though this would likely have to be tested using smaller analysis region areas than that of the 4° x 4° box array used for this analysis.

4.4.2 Temporal Behavior of Onset

As onset varies spatially from year-to-year, the same is true for the temporal characteristics of onset. Figure 59 presents the IPF rainfall time series for each year of the analysis, in addition to the seven-year average IPF time series, averaged for all boxes. These time series present a means of comparing the domain-averaged onset from year-to-year, in addition to comparing a specific year to the seven-year domain-average onset date. In general, the same general trends occur from year-to-year. For each year, there is a gradual increase in IPF rainfall that represents the transition from winter precipitation regimes to those of the summer season. Based on the below time series, IPF precipitation increase typically begins in mid-to-late April (pentads 22-24) and continues into the summer months. Looking at all years except 2008 and 2009, there is only a three-pentad difference in the earliest (pentad 30 in 2010) and latest occurrences (pentad 33 in 2011) of onset. When comparing the yearly onset results to that of the seven-year average, there is a maximum of two pentads difference in onset date. This maximum difference again is due to the year 2011, with a domain-average onset date two pentads later (pentad 33) than that of the seven-year average (pentad 31) (Figure 60). In the seven years that do not include 2008 and 2009, domain-average IPF onset occurs between pentads 30 and 33 (May 26-June 14). As such, IPF onset occurs in late May-to-mid June for these analysis years.

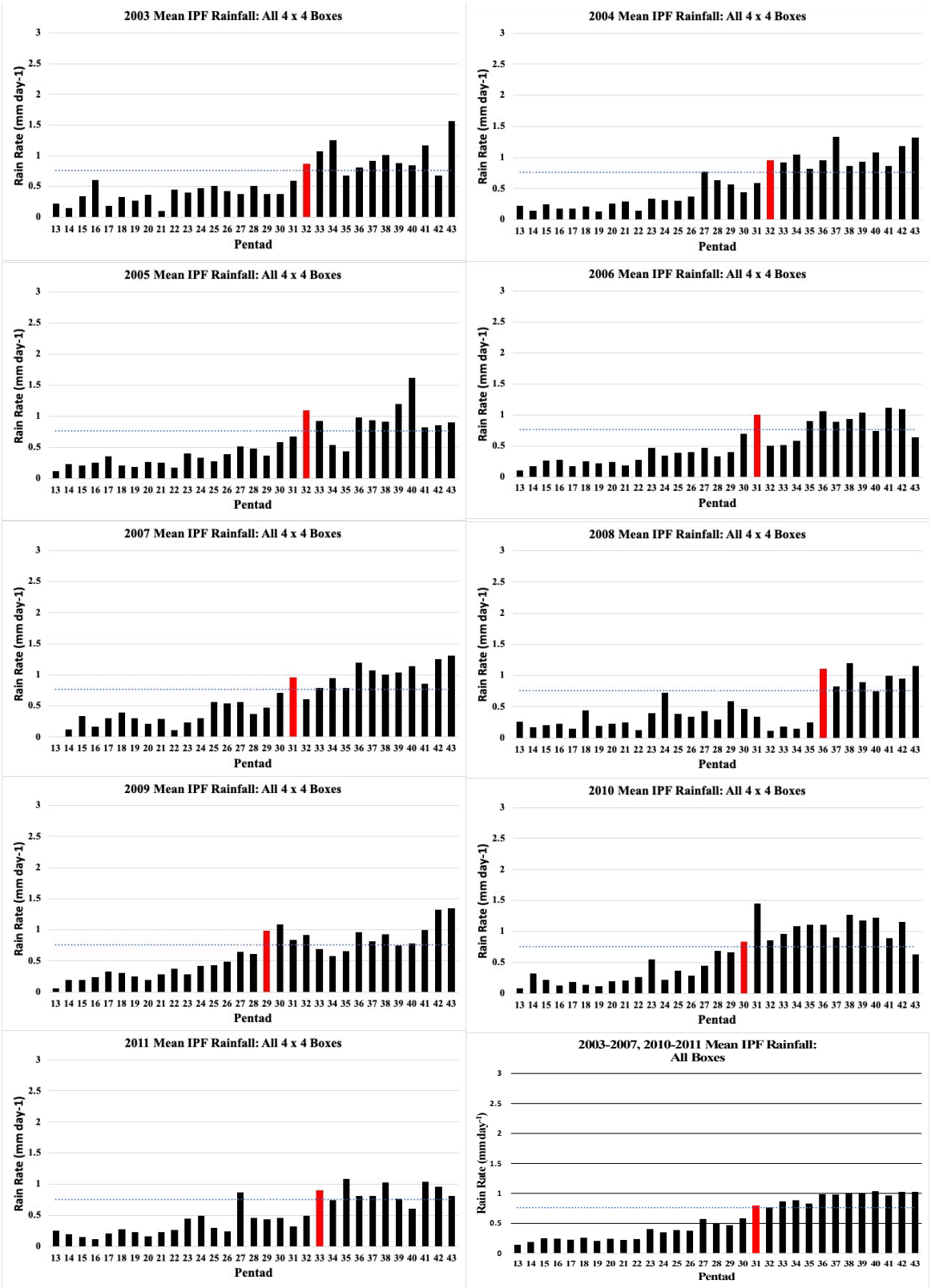


Figure 59: IPF rainfall time series for: (1) each year of the analysis, and (2) the seven-year average. Each time series accounts for the average IPF rainfall of all boxes during each respective pentad.

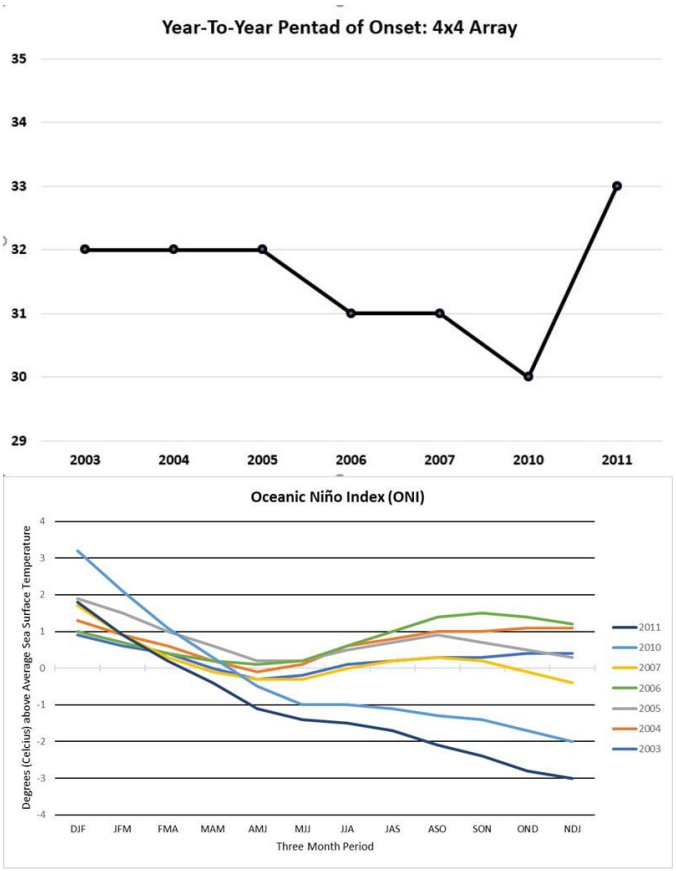


Figure 60: Line graphs showing the change in pentad of onset from year-to-year and the yearly changes in ONI, a metric used to examine ENSO.

4.5 Onset Threshold Sensitivity

As previously mentioned, onset sensitivity was also tested with respect to changes to the threshold rainfall value used to determine the pentad of onset for the seven-year average. The first threshold, used in the analysis described in prior sections, is 0.7638 mm day^{-1} and comes from previous studies of onset for the SE US (Rickenbach et al., 2015, Rickenbach et al., 2020). The second threshold used in the sensitivity analysis is the mean rainfall rate for the entire domain, 0.5017 mm day^{-1}. The idea behind using two different thresholds is a result of two factors: (1) to observe how onset characteristics compare and differ when looking at both thresholds, and (2) to have a means of comparison to previous literature while also exploring onset characteristics with an original threshold. An onset map and IPF rainfall time series (Figure

62), like those in Figures 21 and 22, were created to show onset analysis results for the 4° x 4° box array domain using the new threshold to allow for direct comparison.

Using the new onset threshold rain rate of 0.5017 mm day^{-1}, onset occurs in 33 of the 38 boxes, up from 18 of 38 boxes when 0.7638 mm day^{-1} is used as the onset determining threshold. In both scenarios, the earliest occurrence of onset is at box 38 in south Florida. However, onset at box 38 occurs three pentads earlier (pentad 21) when using the smaller threshold than with the original (pentad 24). Across the entire domain, onset only exceeds pentad 30 four times out of a possible 33 onset occurrences using the smaller threshold. Conversely, onset exceeds pentad 30 11 times out of a possible 18 onset occurrences in the original onset analysis. Further, there is a six-pentad difference in domain-average time of onset, with average onset occurring at pentad 27 with the smaller threshold and at pentad 33 in the original onset analysis. As one may expect in this case, decreasing the threshold rain rate required to determine onset date leads to two things: 1) a decrease in both box-average and box-specific onset timing, and (2) an increase in the number of boxes where onset occurs with respect to the criteria for onset determination. A great

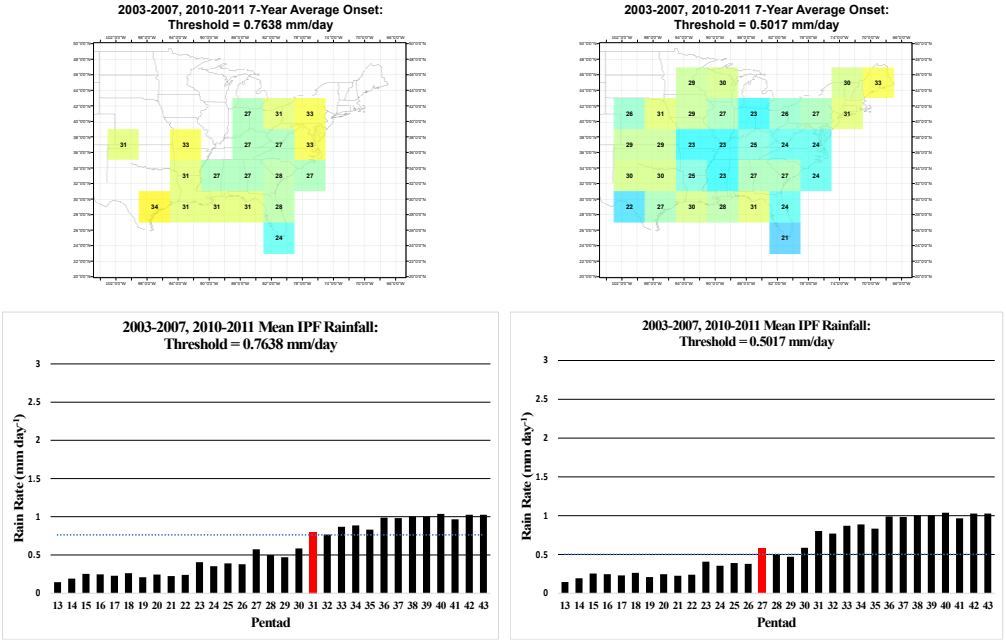


Figure 61: Comparison of seven-year average onset results using two different onset rain rate thresholds.

example as to why these changes occur is shown in boxes 3 and 10 (Figure 62). For boxes 3 and 10, under the original onset threshold of 0.7638 mm day^{-1} onset does not occur at either box. The reason for this is that the box-average pentad rain rate only meets or exceeds the threshold value once at box 3 and three times at box 10. However, when the onset threshold is lowered to 0.5017 mm day^{-1}, the onset criteria are met in both boxes, as the less intense threshold value is now closer to the box-average pentad rain rates of these two boxes.

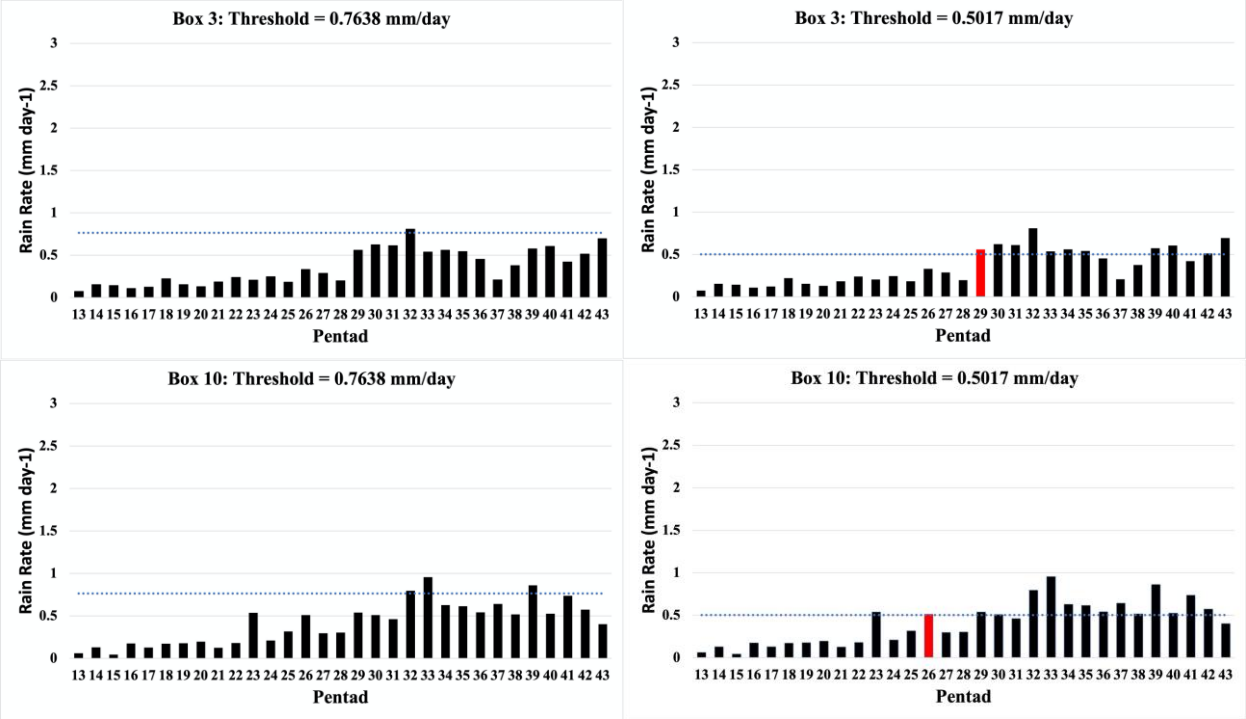


Figure 62: Threshold sensitivity analysis at boxes 3 and 10.

CHAPTER 5: CONCLUSIONS

This project examined the characteristics of the observed springtime onset of the convective season in the SE and central US, with respect to the IPF and MPF convective organization framework to analyze onset. The time and space characteristics of IPF onset occurrence has been studied previously but was limited to a four-year time period and the SE US spatial domain. The time and space characteristics of IPF onset for the central US had not been studied in previous research. An objective method for the determination of IPF onset for the SE and central US was used. The methods used were applied to two different analysis box arrays, in order to test onset sensitivity with respect to changes in spatial analysis region size. Further sensitivity analysis was done using two different onset threshold rain rates, in an effort to test onset sensitivity and compare onset observations as the onset threshold changes in the algorithm.

The date of onset of the IPF convective season for the SE and central US was determined based on the IPF precipitation in each 4° x 4° box. Using the onset criteria, onset was determined for each box within the box array for the years 2003-2011 and for a seven-year average encompassing the years 2003-2007 and 2010-2011. As the results of the seven-year average analysis show, onset in the SE US is observed to progress northward and westward in time, beginning in South Florida in late-April and having westernmost and northernmost extents in eastern Texas and southern New York throughout the months of May and June. Springtime IPF onset is shown to be more clearly defined in the SE US than in the central Plains and northern US, as shown by the spatial pattern of onset as well as the difference in July minus March IPF precipitation. Year-to-year, onset characteristics of the SE US deviate from this trend for a couple of years of the analysis, but onset generally is still a northward and westward expanding phenomenon with respect to both time and space. For the SE US, the occurrence and timing of

onset is more consistent year-to-year in the SE US compared to the expanded northern and western portions of the domain. An initial hypothesis for this onset behavior is that there is a springtime increase in moisture and instability from the seasonal establishment of the NASH that results in IPF onset that is mostly limited to the SE US, though addressing this question is outside of the scope of the thesis. Further, the geographic extent of onset is generally confined to the SE US and typically not present in the central US (Figures 21 and 24).

The areal sensitivity analysis shows a key point in the onset characteristics for both the SE and central US, both related to the analysis regions' corresponding box size. When comparing the 4° x 4° and 2° x 2° onset characteristics, there is greater box-to-box difference in onset date, due primarily the higher resolution of the 2° x 2° boxes. For this reason, the differences in box-to-box onset date for the 2° x 2° tend to be more random in nature. Further, the same systemic temporal and spatial characteristics of the 4° x 4° boxes are also observed for the 2° x 2° boxes.

Missing and invalid data rendered multiple years unusable in the onset analysis. In 2008, about 39% of data was missing for April and about 72% for June. Further, portions of data in May are also missing (?) and are as such invalid. As such, data of high quality is only present in the months of March, June after the 23rd and July for 2008. Onset from box-to-box in 2008 was primarily observed at the end of June, with the exception of one box with onset at the end of March. These onset dates also represent periods where high quality data was available. Since much of April was missing and May was invalid, there was no way to validate or verify the observed onset characteristics of 2008 in this study. 2009 was similar in that both missing and invalid data rendered the onset analysis unusable. While about 8% of April and 31% of June data are completely missing for 2009, it was data quality in the month of May that hurt the analysis

the most. Onset for much of the 2009 onset domain occurs in the month of May, which is not an abnormal result when compared to average or when looking at year-to-year trends in onset. However, because isolated portions of the data for the first half of May were null and data was missing for April and June, there was no way to validate or verify findings related to onset characteristics for the year 2009.

The sensitivity analysis related to the onset threshold rain fall rate reveals two relationships. When the threshold rain rate value is lowered within the algorithm, the average date of onset decreases both box-to-box and for the domain average. Relatedly, decreasing the threshold value yields a greater number of boxes with onset occurrence over the seven-year time averaged period. This is of particular importance for the Central Plains region in the central US, where IPF rainfall does not tend to greatly intensify during the spring-to-summer transitional period. As such, a lower threshold leads to a larger geographical extent of observed onset, particularly for regions of the country where IPF rainfall is less intense from year-to-year. The 0.7638 mm day^{-1} threshold is the more appropriate threshold, as it is based on the average precipitation of the SE US region, where the onset characteristics of the larger domain were evaluated using the onset characteristics of the SE US subregion.

While the results of this research have solidified past findings related to the onset characteristics of the SE US and have provided a basic understanding of onset characteristics for the central US, further expansion of this analysis will be necessary in the future. First, it is of great importance to expand on this work by exploring the meteorological variables at play that influenced the results and observations presented here. In particular, of great importance are the dynamic and thermodynamic characteristics specific to the temporal and spatial domains of this study. It will also be useful to replicate this analysis using a different dataset. This would be

useful in two distinct ways: (1) in an effort to further solidify findings related to the SE US given the expanded temporal period and, (2) to see if the initial findings presented here for the central US hold true when compared to findings in different data sets. Secondly, it may be possible to solidify findings for the years 2008 and 2009. In particular, exploring if machine learning techniques can reasonably predict precipitation and onset characteristics where data is missing in this study. Finally, it is necessary to further expand the temporal period of this analysis as more recent data becomes available. Expanding the temporal period of this study to include more recent data past 2011 would allow for a greater understanding of onset characteristics in both the SE and central US, in addition to expanding on the current onset climatology.

REFERENCES

Barkiđija, S., Fuchs, Ž. 2013. *Precipitation correlation between convective available potential energy, convective inhibition and saturation fraction in middle latitudes.* Atmospheric Research, 124, 170-180.

Carbone, R., Tuttle, J. 2008. *Rainfall Occurrence in the U.S. Warm Season: The Diurnal Cycle.* Journal of Climate, 21, 4132-4146.

Davis, R., Hayden, B., Gay, D., Phillips, W., Jones, G. 1997. *The North Atlantic Subtropical Anticyclone*. Journal of Climate, 10(4), 728-744.

DeRubertis, D. 2006. *Recent Trends in Four Common Stability Indices Derived from U.S. Radiosonde Observations*. Journal of Climate, 19, 309–323.

Diem, J. 2006. *Synoptic-Scale Controls of Summer Precipitation in the Southeastern United States.* Journal of Climate, 19(4), 613-621.

Djurić, D., Ladwig, D. 1983. *Southerly Low-Level Jet in the Winter Cyclones of the Southwestern Great Plains*. Monthly Weather Review, 111, 2275-2281.

Doswell III, C., Markowski, P. 2004. *Is Buoyancy a Relative Quality?* Monthly Weather Review, 132, 853-863.

Feng, Z., Houze, R., Leung, L.R., Song, F., Hardin, J., Wang, J., Gustafson, W., Homeyer, C. 2019. *Spatiotemporal Characteristics and Large-Scale Environments of Mesoscale Convective Systems East of the Rocky Mountains*. Journal of Climate, 32, 7303–7328.

Gates, W. 1960. *Static Stability Measures in the Atmosphere*. Journal of Meteorology, 18, 526-533.

Haberlie, A., Ashley, W. 2018. *A Radar-Based Climatology of Mesoscale Convective Systems in the United States.* Journal of Climate, 32(5), 1591-1606.

Holton, J. 2004. *An Introduction to Dynamic Meteorology*. Academic Press, 535 pp.

Houze, R. 1989. *Observed structure of mesoscale convective systems and implications for large-scale heating*. Quarterly Journal of the Royal Meteorological Society, 115, 425-461.

Houze, R. 1993. *Cloud Dynamics*. Academic Press, 573 pp.

Houze, R. 2004. *Mesoscale Convective Systems.* Review of Geophysics, 42 (4), 1-43.

Houze, R., Lee, W., Bell, M. 2009. *Convective Contribution to the Genesis of Hurricane Ophelia (2005)*. Monthly Weather Review, 137, 2778-2800.

Klimowski, B., Bunkers, M., Hjelmfelt, M., and Covert, J. 2003. *Severe Convective Windstorms over the Northern High Plains of the United States*. Weather and Forecasting, 18, 502 -519.

Kirshbaum, D., Adler, B., Kalthoff, N., Barthlott, C., Serafin, S. 2018. *Moist Orographic Convection: Physical Mechanisms and Links to Surface-Exchange Processes*. Atmosphere 2018, 9, 80.

Koch, S., Ray, C. 1997. *Mesoanalysis of summertime convergence zones in central and eastern North Carolina*. Weather and Forecasting, 12, 56-77.

Kousky, V. 1988. *Pentad Outgoing Longwave Radiation Climatology for the South American Sector.* Revista Brasileira de Meteorologia, 3, 217-231.

Leary, C., Houze, R. 1979. *The Structure and Evolution of Convection in a Tropical Cloud Cluster*. Journal of the Atmospheric Sciences, 36, 437–457.

Leary, L., Ritchie, E. 2010. *The Role of Convection in Determining Tropical Cyclone Genesis in the Eastern North Pacific*. 29th Conference on Hurricanes and Tropical Meteorology, Tucson, AZ, American Meteorological Society, 16D.5, https://ams.confex.com/ams/29Hurricanes/techprogram/paper_169209.htm.

Li, W., Laifang, L., Fu, R., Deng, Y., Wang, H. 2011. *Changes to the North Atlantic Subtropical High and Its Role in the Intensification of Summer Rainfall Variability in the Southeastern United States.* Journal of Climate, 24, 1499-1506.

Luchetti, N., Nieto-Ferreira, R., Rickenbach, T., Nissenbaum, M., and McAuliffe, J. 2017. *Influence of the North Atlantic subtropical high on wet and dry sea-breeze events in North Carolina, United States*. Investigaciones Geográficas, 68, 9-25.

Marengo, J., Liebmann, B., Kousky, V., Filizola, N., Wainer, I. 2001. *Onset and End of the Rainy Season in the Brazilian Amazon Basin*. Journal of Climate, 14, 833–852.

Markham, C. 1970. *Seasonality of Precipitation in the United States*. Annals of the Association of American Geographers, 60(3), 593-597.

Miller, P., Mote, T. 2017. *A Climatology of Weakly Forced and Pulse Thunderstorms in the Southeast United States.* Journal of Applied Meteorology and Climatology, 56, 3017 -3033.

Moncrieff, M. 2010. *The Multiscale Organization of Moist Convection and the Intersection of Weather and Climate*. Geophysical Monograph Series, No.189, American Geophysical Union, 3-26.

Nesbitt, S., Cifelli, R., Rutledge, S. 2006. *Storm morphology and rainfall characteristics of TRMM precipitation features*. Monthly Weather Review, 134, 2702-2721.

Nieto-Ferreira, R., Hall, L., Rickenbach, T. 2013. *A climatology of the structure, evolution, and propagation of midlatitude cyclones in the southeast United States*. Journal of Climate, 26, 8406-8421.

Nieto-Ferreira, R., Rickenbach, T. 2011. *Regionality of monsoon onset in South America:a three-stage conceptual model.* International Journal of Climatology, 31, 1309-1321.

Nieto-Ferreira R., and Rickenbach, T., 2021: *Mechanisms for springtime onset of isolated convection precipitation across the Southeastern United States*. Atmosphere, 12, 213.

Nieto-Ferreira R., and Rickenbach, T., 2020: *Effects of the North Atlantic Subtropical High on summertime precipitation organization in the southeast United States.* International Journal of Climatology. 40 (14), 5987-6001.

Nieto-Ferreira, R. and Rickenbach, T., 2021. Mechanisms for springtime onset of isolated precipitation across the Southeastern United States. *Atmosphere*, 12(2), 213.

Peppier, R. 1988. *A review of static stability indices and related thermodynamic parameters.* Illinois State Water Survey Misc. Publ. 104, 87 pp. [Available from Illinois State Water Survey, Champaign, IL 61820.].

Prat, O., Nelson, B. 2013. *Precipitation Contribution of Tropical Cyclones in the Southeastern United States from 1998 to 2009 Using TRMM Satellite Data*. Journal of Climate, 26, 1047-1062.

Prat, O., Nelson, B. 2014. *Characteristics of annual, seasonal, and diurnal precipitation derived from long-term remotely sensed data*. Atmospheric Research, 144, 4-20.

Prat, O., Nelson, B. 2015. *Evaluation of precipitation estimates over CONUS derived from satellite, radar, and rain gauge data sets at daily to annual scales (2002–2012).* Hydrology and Earth System Sciences, 19, 2037-2056.

Rickenbach, T. M., Nieto-Ferreira, R., Zarzar, C., Nelson, B. 2015. *A seasonal and diurnal climatology of precipitation organization in the southeastern United States*. Quarterly Journal of the Royal Meteorological Society, 141 (690), 1938-1956.

Rickenbach, T.M., Nieto Ferreira, R., & Wells, H. 2020. *Springtime onset of isolated convection precipitation across the southeastern United States: Framework and regional evolution*. Monthly Weather Review.

Seager, R., Murtugudde, R., Naik, N., Clement, A., Gordon, N., Miller, J. 2003. *Air-Sea Interaction and the Seasonal Cycle of the Subtropical Anticyclones*. Journal of Climate, 16(12), 1948-1966.

Sherburn, K., Parker, M. 2014. *Climatology and Ingredients of Significant Severe Convection in High-Shear, Low-CAPE Environments*. Weather and Forecasting, 29, 854–877,

Smith, R. 2000. *The Role of Cumulus Convection in Hurricanes and its Representation in Hurricane Models*. Review of Geophysics, 38(4), 465-489.

Wallace, J. 1975. *Diurnal Variations in Precipitation and Thunderstorm Frequency over the Conterminous United States*. Monthly Weather Review, 103, 406–419.

Wang, B., Ding, Q., Joseph, P. 2009. *Objective Definition of the Indian Summer Monsoon Onset*. Journal of Climate, 22, 3303–3316.

Wilson, J., Schreiber, W. 1986. *Initiation of Convective Storms at Radar-Observed Boundary Layer Convergence Lines.* Monthly Weather Review, 114, 2516-2536.

Zhang, J., Howard, K., Langston, C., Vasiloff, S., Kaney, B., Arthur, A., Van Cooten, S., Kelleher., K,Kitzmiller, D., Ding, F., Seo, D-J., Wells, E., Dempsey, C. 2011. *National Mosaic and multi-sensor QPE (NMQ) system: Description, results, and future plans.* Bulletin of the American Meteorological Society, 92, 1321–1338.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., Grams, H., Wang, Y., Cocks, S., Martinaitis, S., Arthur, A., Cooper, K., Brogden, J., Kitzmiller, D. 2016. *Multi-Radar Multi-Sensor (MRMS) Quantitative Precipitation Estimation: Initial Operating Capabilities.* Bulletin of the American Meteorological Society, 97, 621–638.

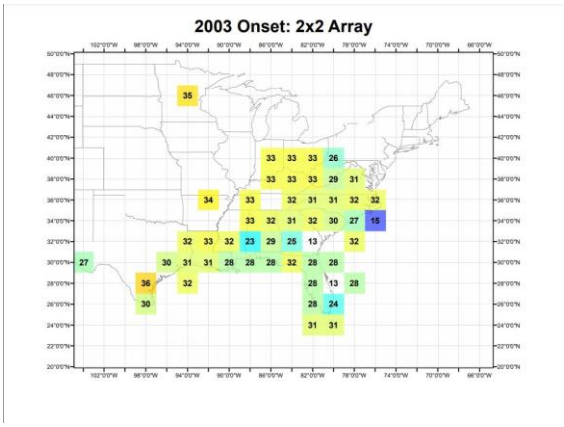
APPENDIX A: PYTHON CODE

1. Python script used for the onset analysis algorithm

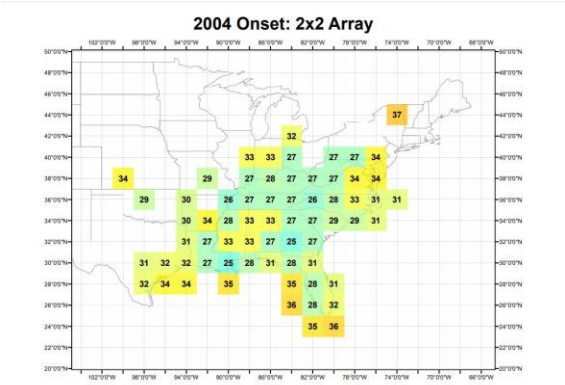
```
### Christopher M. Jarrett ###
### Created: 8/18/2020, 16:14 ###
### This python script is used to identify rainfall rates that are greater than or equal to a threshold value ###
### Boolean values (True/False) indicate whether a given pentad's rainfall rate equals or exceeds the threshold ###
#
### Import the pandas package ###
import pandas as pd
#
### Define data to be read into pandas' data frame ###
data=pd.read_csv('Avg2.csv')
#
### Define and print the pandas' dataframe ###
df = pd.DataFrame(data=data)
df
#
### Create and print a new dataframe with boolean outputs using the threshold rainrate ###
new_df = (df >= 0.7638)
new_df
#
### Export the new dataframe to a csv file containing the boolean outputs, to be manipulated in excel manually ###
new_df.to_csv('NewAvg2x2.csv')
#
exit()
```

APPENDIX B: INDIVIDUAL 2° x 2° FIGURES

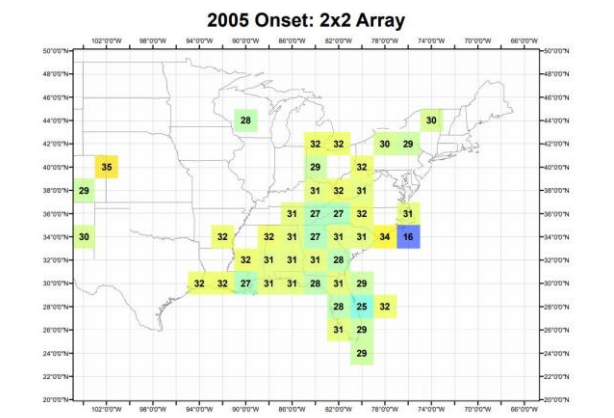
1. 2003



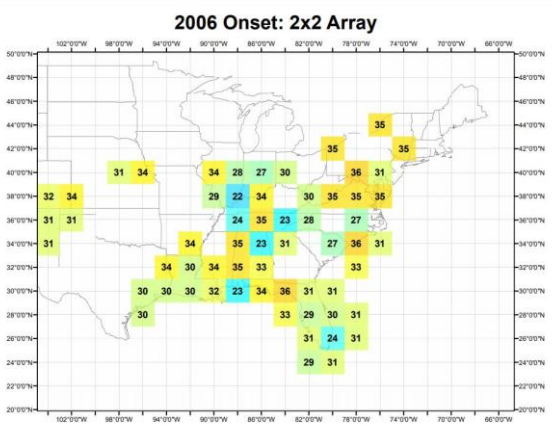
2. 2004



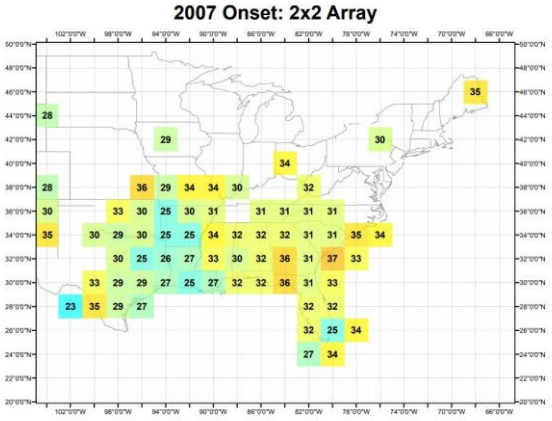
3. 2005



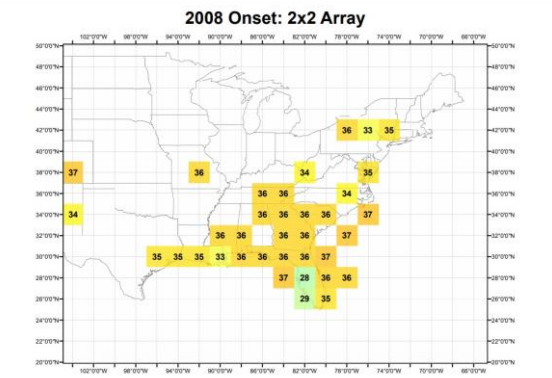
4. 2006



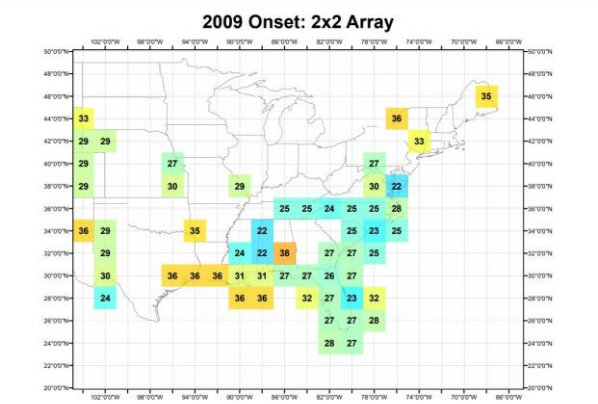
5. 2007



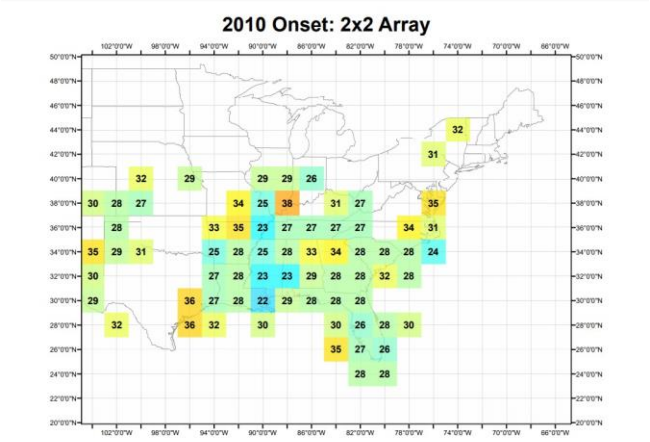
6. 2008



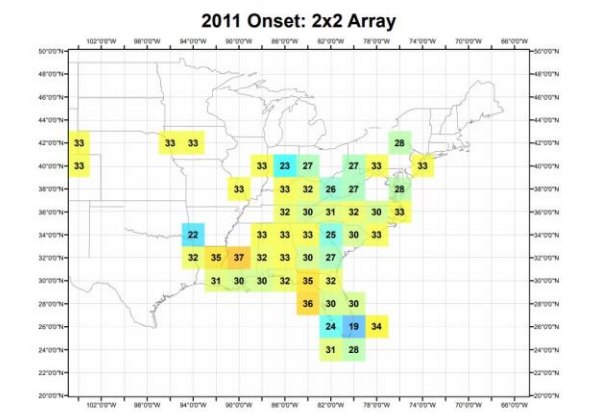
7. 2009



8. 2010



9. 2011



10. 2003-2007, 2010-2011 seven-year average